

# **8th AIUCD CONFERENCE 2019**

Pedagogy, Teaching, and Research  
in the Age of Digital Humanities

Udine, Italy, January 22-25, 2019

**Ottavo Convegno Annuale**  
**8th Annual Conference**  
**AIUCD 2019**  
**Udine, 23 – 25 gennaio 2019**

## **Book of Abstracts**

**Didattica e ricerca**  
**al tempo delle Digital Humanities**

***Teaching and research***  
***in Digital Humanities' era***

edited by Stefano Allegrezza

ASSOCIAZIONE PER  
L'INFORMATICA UMANISTICA  
E LA CULTURA DIGITALE



**2019**

**ISBN: 978-88-942535-3-5**

Copyright © 2019

ASSOCIAZIONE PER  
L'INFORMATICA UMANISTICA  
E LA CULTURA DIGITALE



Associazione per l'Informatica Umanistica e la Cultura Digitale



Copyright of each individual chapter is maintained by the authors.

This work is licensed under a Creative Commons Attribution Share-Alike 4.0 International license (CC-BY-SA 4.0). This license allows you to share, copy, distribute and transmit the text; to adapt the text and to make commercial use of the text providing attribution is made to the authors (but not in any way that suggests that they endorse you or your use of the work). Attribution should include the following information:

Stefano Allegrezza (ed.), AIUCD 2019 - *Book of Abstracts*, Udine 2019.

Cover image has been created by Marco De Anna – Università degli Studi di Udine.

If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

All links have been visited on 5th January 2019, unless otherwise indicated.

Every effort has been made to identify and contact copyright holders and any omission or error will be corrected if notified to the editor: [stefano.allegrezza@unibo.it](mailto:stefano.allegrezza@unibo.it)

# Organizzazione

Gli abstract pubblicati in questo volume hanno ottenuto il parere favorevole da parte di valutatori esperti della materia, attraverso un processo di revisione anonima mediante double-blind peer review sotto la responsabilità del Comitato Scientifico di AIUCD 2019.

Il programma della conferenza AIUCD 2019 è disponibile online all'indirizzo <http://www.aiucd2019.uniud.it/>.

*All abstracts published in this volume have received favourable reviews by experts in the field of DH, through an anonymous double peer review process under the responsibility of the AIUCD 2019 Scientific Committee. The AIUCD 2019 conference program is available online at <http://www.aiucd2019.uniud.it/>.*

## International Programme Committee

Stefano Allegrezza (Chair) – Università di Bologna  
Valentina Bartalesi – CNR Pisa  
Marina Buzzoni – Università Ca' Foscari Venezia  
Alberto Campagnolo – Università di Udine  
Emanuela Colombi (Co-Chair) – Università di Udine  
Rocco Coronato – Università di Padova  
Giorgio Maria Di Nunzio – Università di Padova  
Tiziana Mancinelli – Universität zu Köln  
Tommaso Mazzoli – Università di Udine  
Paolo Monella – Università di Palermo  
Stefano Ondelli – Università di Trieste  
Francesco Pitassio – Università di Udine  
Simone Reborà – Georg-August-Universität Göttingen  
Daria Spampinato (ex officio) – CNR Catania  
Rachele Sprugnoli – Fondazione Bruno Kessler Trento  
Francesca Tomasi – Università di Bologna  
Arjuna Tuzzi – Università di Padova

## Local Organization Committee

Emanuela Colombi (Co-Chair) – Università di Udine  
Tommaso Mazzoli – Università di Udine  
Francesco Pitassio – Università di Udine

# Indice

<b>Prefazione .....</b>	<b>8</b>
<b>AIUCD 2019 Panels.....</b>	<b>11</b>
<b>Models as forms, models as concepts</b>	
Øyvind Eide.....	12
<b>Finding common ground between text, maps, and tables for quantitative and qualitative research</b>	
Marieke Van Erp .....	13
<b>Are the Digital Humanities ‘language insensitive’? Connecting debates about Modern Languages, global cultural representation in DH and the international classroom</b>	
Paul Spence .....	14
<b>AIUCD2019 Panels.....</b>	<b>15</b>
<b>Metodologie informatiche per le discipline umanistiche: stato dell’arte, criticità e proposte per una classe di laurea magistrale</b>	
Francesca Tomasi, Alessandro Lenci, Anna Rovella, Franco Tommasi .....	16
<b>Filosofia e Digitale?/Digital Philosophy?</b>	
Luca Bandirali .....	21
<b>L'uso degli applicativi HGIS-WebGIS nella ricerca geo-storica</b>	
Tiago Luis Gil, Massimiliano Grava, Nicola Gabelliere, Vinicius Maluly .....	29
<b>WORKSHOPS.....</b>	<b>35</b>
<b>Visualizing the Italian Literary Canon through Distant Reading</b>	
Tiziana Mancinelli, Simone Rebora, Rachele Sprugnoli .....	36
<b>LONG PAPERS.....</b>	<b>39</b>
<b>Per una filologia (digitale) degli scartafacci</b>	
Elena Pierazzo, Alessia Marini .....	40
<b>Can we do better than computers? Teaching Computational Thinking to Digital Humanists</b>	
Marilena Daquino, Silvio Peroni, Francesca Tomasi.....	43
<b>Developing digital research workflows among undergraduates</b>	
Mike Cosgrave .....	47
<b>L’ecosistema digitale per la cultura del piemonte. Strumenti e modelli per la condivisione della conoscenza dei beni culturali archivistici, librari e museali</b>	
Dimitri Brunetti .....	54

<b>SPARSAR recites Shakespeare’s Sonnets – and Coping with Early Modern English variants</b>	
Rodolfo Delmonte .....	59
<b>“Digital Autoethnography” and “Connective Intelligence”: Teaching and (Re)Thinking (About) Today’s Techno-Society</b>	
Stefano Calzati .....	63
<b>Oltre la galassia delle Digital Humanities: per la costituzione di una disciplina di Informatica Umanistica</b>	
Fabio Ciotti.....	67
<b>Per un archivio digitale dell’Università Castrense di San Giorgio di Nogaro: saperi, pratiche, immagini durante la Grande guerra</b>	
Dario De Santis .....	73
<b>Funzione dell’edizione critica digitale nel contesto degli studi universitari. Questioni di scelte didattiche: studenti codificatori o studenti esecutori?</b>	
Luciano Longo, Alba Castello, Natalia Librizzi, Ugo La Mantia.....	78
<b>The Digital Repository Service of the National Documentation Centre in Greece: a model for Digital Humanities data management and representation</b>	
Katerina Bartzi, Nikos Vasilogamvrakis, Elena Lagoudi, Despina Hardouveli, Evi Sachini.....	83
<b>Didattica universitaria e ambienti digitali</b>	
Chiara Panciroli, Anita Macauda .....	91
<b>The Use of Blockchain for Digital Archives: Challenges and Perspectives</b>	
Clara Bacciu, Angelica Lo Duca Andrea Marchetti .....	99
<b>L’Orientale Foundation: un programma per l’introduzione delle nuove tecnologie nei curricula umanistici</b>	
Johanna Monti, Valeria Caruso .....	105
<b>Towards a Computational Stylistics of Book Reviews</b>	
Massimo Salgaro, Simone Rebora .....	110
<b>Lingue antiche e Digital Humanities: annotazione digitale dei testi tra ricerca e didattica</b>	
G. Mugelli, G. Re, A. Taddei .....	117
<b>Integrazione di dimensioni narrative e modalità immersive di esplorazione interattiva in visite educative museali</b>	
Giovanni Luca Dierna, Alberto Machì.....	125
<b>Narrazioni e frammenti, mosaici e algoritmi. La costruzione digitale del racconto tra spazialità e temporalità</b>	
Federico Meschini .....	133
<b>Il Trattato di Scienza Universal di Vivaldo Belcalzer: fra tradizione e segmentazione testuale. Prove di dialogo ecdotico</b>	
Rosa Casapullo, Luciano Longo .....	135
<b>Diogene alla ricerca dell’uomo contemporaneo: le Digital Humanities “lucerna” per riconoscere il Digital Cultural Heritage. La riflessione DiCultHer-AIUCD</b>	
Nicola Barbuti, Fabio Ciotti, G. De Felice.....	140

<b>Creating Digital Culture by digitizing Cultural Heritage: the Crowddreaming living lab method</b>	
N. Barbuti, Giuliano De Felice, Annalisa Di Zanni, P. Russo, A. Valentini .....	145
<b>On the Use of Terminological Records in Specialised Translation</b>	
Federica Vezzani, Giorgio Maria Di Nunzio .....	151
<b>EpiCUM. Un museo epigrafico digitale per visitatori e specialisti</b>	
Salvatore Cristofaro, Daria Spampinato.....	157
<b>Digital Humanities and “Niche” Research Fields: The Case of Ancient Arabian Epigraphy</b>	
Irene Rossi.....	164
<b>Gli archivi personali nell’era digitale: stato dell’arte, criticità e prospettive</b>	
Stefano Allegrezza .....	170
<b>L’utilizzo dei database da parte degli storici: storiografia e dibattito attuale</b>	
Tiago Luis Gil .....	177
<b>SHORT PAPERS .....</b>	<b>182</b>
<b>Analizzare, modellizzare e codificare i pareri di lettura editoriali</b>	
Laura Antonietti .....	183
<b>eLearning the URLCoFi - Digital Didactics for Humanists</b>	
Armin Hoenen .....	187
<b>14-18 Documenti e immagini della Grande Guerra</b>	
Elisa Sciotti, Francesco Gandolfi .....	190
<b>Il Laboratorio di Epigrafia Greca dell’Università Ca’ Foscari. Una fucina didattica per l’epigrafia greca</b>	
Eloisa Paganoni, Stefania De Vido, Claudia Antonetti.....	193
<b>Complessità della codifica ed ergonomia strumentale nel contesto XML-TEI: dove siamo? (Bilancio a partire da un nuovo progetto di edizione digitale medievale)</b>	
Marta Materni.....	196
<b>“Osservate, leggete con me” Risorse LOD per la storia del melodramma: una prospettiva funzionale di rappresentazione</b>	
Paolo Bonora, Angelo Pompilio.....	202
<b>Le trascrizioni dei sogni dei sopravvissuti di Auschwitz: analisi linguistica e tematica secondo un modello Xml-Tei</b>	
Carlo Chiaramonte, Francesco Figoli, Simone Mastrocesare, Daniele Silvi, Daniele Tosco, Marco Zanasi .....	208
<b>Constructing Narratives Using NBVT: A Case Study</b>	
Valentina Bartalesi, Daniele Metilli, Carlo Meghini .....	218
<b>Progettare, realizzare e promuovere contenuti culturali digitali. Un esempio di didattica integrata: il progetto “Filosofia &amp; Migrazioni”</b>	
Maria Eugenia Cadeddu, Francesca Gambetti, Cristina Marras, Ada Russo.....	223
<b>Serious Games as immersive educational tools: a model for the Pathways of Romanticism (Porto, Portugal)</b>	
João Victor Camara, Alice Lucas Semedo.....	227

<b>A Corpus Linguistic Approach to Pronouns in a Work of Fiction: The Narrative Voices in Conrad's <i>The Shadow Line</i></b> .....	<b>233</b>
Giuseppina Balossi.....	233
<b>Corpora in classe – il caso della lingua serba per italofofi</b>	
Olja Perišić Arsić .....	244
<b>Un'indagine sulle competenze digitali di studenti di facoltà umanistiche nativi digitali di ultima generazione</b>	
Floriana C. Sciumbata.....	247
<b>Crowdsourcing Peer Review in the Digital Humanities?</b>	
Michael Soprano, Stefano Mizzaro .....	251
<b>Le rappresentazioni digitali e le edizioni di documenti medievali online per la ricerca e la didattica della Diplomatica e della Storia della documentazione. Un caso concreto</b>	
Antonella Ambrosio, Maria Rosaria Falcone, Vera Isabell Schwarz-Ricci, Georg Vogeler .....	255
<b>La monografia digitale: una riflessione sulle implicazioni dell'innovazione testuale</b>	
Alessandra Di Tella.....	260
<b>Manzoni digitale: un laboratorio tra didattica e ricerca</b>	
Beatrice Nava .....	264
<b>Storia dell'informatica: metodi e strumenti per raccontarla da informatici</b>	
Viola Bongini, Giovanni Antonio Cignoni, Emanuele Lenzi, Nicolò Fratelli.....	269
<b>From collaborative transcription to interdisciplinary education: the postcards of the Great War case</b>	
Enrica Salvatori, Federico Boschetti, Angelo Mario Del Grosso .....	274
<b>Mapping the Persecution of Trentino People Deported to the 3rd Reich Camps</b>	
Rachele Sprugnoli, Alessio Palmero Aprosio, Giovanni Moretti, Sara Tonelli .....	279
<b>Text in metamorphosis: new aesthetics of digital humanities</b>	
Herbert Natta .....	285
<b>Un progetto per la consultazione on-line degli archivi del restauro (1850-1915)</b>	
Gabriella Guarisco, Simonetta Ciranna, Chiara Devoti, Marco Felli, Elena Fioretto, Vittorio Foramitti, Rossana Gabaglio, Nora Lombardini, Alessandra Maniaci, Monica Naretto, Giuseppina Perusini, Elisa Piolatto, Martina Visentin.....	289
<b>Strumenti di Laboratorio. Il Laboratorio Informatico per la Documentazione Storico Artistica del Dipartimento di Studi Umanistici e del Patrimonio Culturale dell'Università di Udine tra ricerca e didattica</b>	
Martina Visentin.....	296
<b>Fondazione Memofonte: l'informatica «aiuta a razionalizzare la ricerca»</b>	
Martina Nastasi, Donata Levi.....	299
<b>The role of research infrastructures in the area of DH education, training and competence building</b>	
Tanja Wissik.....	301
<b>Per un'edizione <i>online</i> dell'Epistolario di Alcide De Gasperi. Criteri di digitalizzazione, schedatura, regestazione ed edizione di lettere del Novecento</b>	
Stefano Malfatti.....	312



<b>POSTER PAPERS .....</b>	<b>319</b>
<b>Can the “digital” reinforce a museum’s mission - what information technologies to introduce into the planned exhibition?</b>	
J. Valach, M. Eisler, P. Štefcová.....	320
<b>Discovering Research Themes in Scientific Research: from Keyphrase Extraction to Co-occurrence Networks</b>	
Rachele Sprugnoli, Giovanni Moretti .....	323
<b>Hyper Spectral Imaging and the Herlufsholm Special Collection</b>	
Jakob Povl Holck, Mogens Kragst Jensen, Kamilla Jensen Husen, Anne Helle Jespersen .....	328
<b>La <i>gamification</i> nella storia dell’arte e nella museologia: impatto sulla metodologia</b>	
Ana Knežević .....	329
<b>LiLa: Linking Latin. Building a Knowledge Base of Linguistic Resources for Latin</b>	
Marco Passarotti, Flavio M. Cecchini, Greta Franzini, Eleonora Litta, Francesco Mambrini, Paolo Ruffolo .....	333
<b>Misurazione degli effetti organizzativi dei progetti digitali nelle reti documentali complesse</b>	
Brizio Tommasi.....	336
<b>Strolling around the dawn of Digital Humanities</b>	
Gabriella Pardelli, Sara Goggi, Federico Boschetti .....	344
<b>Time machines for Online Services: An Evaluation of a New Interface to Visualize Knowledge Over Time</b>	
James Blustein, Nicola Raffaele Di Matteo .....	348
<b>Uno sguardo nei depositi: l’esposizione multimediale dei reperti ceramici di Castelnovo del Friuli</b>	
Irene Sarcinelli .....	351
<b>Vita e i libri di guarnerio d’artegna: un percorso didattico nel digital storytelling laboratory</b>	
Matteo Venier.....	355
<b>A Catalog of Web App for Smart Teaching</b>	
Marco Corbato, Antonina Dattolo.....	359

# Prefazione

Stefano Allegrezza, Università degli Studi di Bologna  
Chair del Comitato di Programma

L'ottava edizione del Convegno nazionale dell'Associazione per l'Informatica Umanistica e la Cultura Digitale, collocandosi nel solco delle precedenti edizioni, si pone l'obiettivo di costituire un punto di incontro tra studiosi delle diverse discipline che - a vario titolo - fanno capo alle cosiddette Digital Humanities: umanisti digitali, informatici, linguisti, storici, archeologi, musicologi, filologici e non solo. Ogni anno la conferenza si propone inoltre di dialogare con altre associazioni del settore per diventare un momento di confronto sempre più ricco e ampio tra gli studiosi.

Questa ottava edizione si svolge per la prima volta in una città del nord-est d'Italia, a Udine<sup>1</sup>, e questo è sintomo da una parte della crescente attenzione che l'Associazione riscuote e di una sua presenza più omogeneamente diffusa sul territorio nazionale, dall'altra dell'interesse dell'Ateneo friulano, ed in particolare del Dipartimento di Studi umanistici e del Patrimonio culturale (il DIUM), verso le tematiche delle Digital Humanities.

La preparazione di un convegno di questa portata parte da molto lontano, ed esattamente dal mese di dicembre 2017 quando il Direttore del DIUM, rispondendo ad una specifica Call dell'Associazione, inviò la propria candidatura ad ospitare ad Udine l'edizione 2019 del convegno AIUCD.

In realtà, se volessimo risalire alle veri "origini" del convegno dovremmo andare ancora più indietro nel tempo, ed esattamente all'anno 2016 quando nel Piano Strategico Dipartimentale il DIUM aveva previsto come uno degli obiettivi principali la costituzione di un Centro di ricerca di primo livello per le digital humanities, cioè "per l'applicazione delle tecnologie digitali allo studio delle discipline umanistiche e del patrimonio culturale, nonché per la riflessione sul loro uso, anche in termini di comunicazione e diffusione della ricerca, soprattutto in relazione allo sviluppo del Web e dei social media".

Quindi, una vocazione alle digital humanities già presente nel Dipartimento e che necessitava di una sua manifestazione di alto livello.

La proposta fu sottoposta al vaglio del comitato direttivo dell'Associazione insieme alle altre pervenute; e il 22 gennaio 2018, esattamente un anno fa, giunse la risposta del Presidente dell'Associazione che assegnava all'Università di Udine l'organizzazione del convegno; la decisione si era basata "sulla qualità della proposta presentata ma anche sulla scelta strategica fatta dal Direttivo di favorire la massima circolazione della conferenza nel territorio nazionale".

Da quel momento si è messa in moto la macchina organizzativa che ha lavorato alacremente in questi dodicimila mesi e con un impegno via via maggiore nel tempo (con una fortissima accelerazione negli ultimi due-tre mesi) e che ha condotto alla realizzazione del convegno al quale diamo ufficialmente inizio. Speriamo ovviamente di soddisfare le aspettative dell'uditorio e degli studiosi qui presenti.

---

<sup>1</sup> Ricordo che la prima edizione si è tenuta a Firenze nel 2012, la seconda a Padova nel 2013, poi a Bologna nel 2014, quindi a Torino nel 2015, a Venezia nel 2016, a Roma nel 2017 e, infine, a Bari lo scorso anno.

Il **tema principale** di questa ottava edizione è la “Didattica e ricerca al tempo delle Digital Humanities”. Si tratta di un ambito vasto che accomuna pratiche e settori di studio interdisciplinari e multidisciplinari e che riguarda i mutamenti della didattica e della ricerca nell'era digitale.

Il convegno vuole avviare una riflessione sulle nuove possibilità rappresentate dal digitale per la didattica e la ricerca scientifica e rispondere a domande del tipo: come il digitale cambierà l'insegnamento delle materie umanistiche? Quale contributo può dare la critica culturale di stampo umanistico alla rivoluzione digitale? Quale relazione sussiste con il piano di digitalizzazione dell'Università lanciato dal Ministero?

Si intende però riflettere anche sulle Digital Humanities come nuovo ambito disciplinare, facendo emergere nuovi interrogativi: come formare la nuova figura del digital humanist? Quali saperi sono chiamati a definire le Digital Humanities come ambito di studio, di ricerca e di formazione? Come riconoscere, classificare, descrivere e valutare la ricerca nel settore delle Digital Humanities?

Nonostante l'argomento principale del convegno - che poteva a prima vista sembrare meno “tradizionale” del consueto - sono pervenuti **75 contributi** presentati da **172 tra autori e co-autori**, non solo italiani ma anche europei ed extraeuropei.

I 75 contributi sono così distribuiti: 31 long papers, 33 short papers, 3 panel, 2 workshop, 6 poster.

La valutazione delle proposte è stata condotta mediante **double-blind peer review** da studiosi competenti in discipline umanistiche, informatiche e di Digital Humanities. Ciascun contributo è stato sottoposto di norma a 3 valutazioni. In alcuni casi (ad esempio nel caso di contributi border-line o di valutazioni contrastanti tra loro) il numero di revisioni è arrivato fino a 6. In totale le valutazioni assegnate sono state ben **258**, ovvero una media di 3,44 revisioni per ciascun contributo. Per svolgere tali valutazioni sono stati coinvolti **122 valutatori** provenienti non solo dall'Italia ma anche da numerosi Stati esteri (33 su 122, ovvero il 27%)<sup>2</sup>.

Dopo il processo di review, dei 75 contributi presentati, 59 sono stati accettati nella stessa tipologia in cui erano stati presentati (long papers, short papers), 7 sono stati accettati secondo una tipologia differente (poster), mentre 9 (pari al 12,2% del totale) non hanno superato la valutazione. Quindi, in totale sono stati accettati 66 contributi su 75, una percentuale di accettazione molto alta (88%) a testimonianza della qualità delle proposte presentate.

Alla fine il processo di revisione ha portato ad avere 26 long papers, 28 short papers, 10 posters, 3 panel, 1 preconference workshop.

A questi si devono aggiungere la **tavola rotonda** sul riconoscimento accademico delle Digital Humanities e le relazioni presentate dal keynote speaker e da tre invited speakers.

---

<sup>2</sup> La maggior parte dei valutatori appartiene alle aree CUN (Consiglio Nazionale Nazionale) 10 - Scienze dell'antichità, filologico-letterarie e storico-artistiche e 11 - Scienze storiche, filosofiche, pedagogiche e psicologiche, ambiti in cui l'umanistica digitale è storicamente più forte e che esprimono ancora oggi una forte prevalenza nell'Associazione. Si registra tuttavia anche la presenza di valutatori che rientrano in altre aree, come l'area 8 - Ingegneria civile ed architettura e 12 - Scienze giuridiche e nell'area delle Digital Humanities (attribuita a studiosi stranieri in quanto non ancora riconosciuta dal sistema italiano), aree 1 e 9 delle Scienze informatiche e ingegneristiche.

Le proposte presentate coprono moltissimi ambiti delle ricerche di Informatica Umanistica, facendone emergere al meglio l'aspetto caratteristico e ineliminabile dell'interdisciplinarietà.

In conclusione, prima di dare inizio ai lavori, permettetemi di ringraziare il **Presidente dell'associazione AIUCD Fabio Ciotti** che è stato per me un riferimento costante, una vera e propria "stella polare" di fronte alle varie difficoltà che sono emerse durante le varie fasi di organizzazione dei lavori; il **chair della precedente edizione del convegno**, Daria Spampinato, che mi ha fornito un continuo supporto anche operativo (vedasi utilizzo di ConfTool), il **Direttore del Dipartimento di Studi umanistici e del patrimonio culturale**, Andrea Zannini, che è stato la vera anima propulsiva di questo evento.

Un caloroso ringraziamento va anche a tutti i membri del **Comitato di Programma** con cui ho condiviso, nello spirito della massima apertura, le diverse decisioni; ai valutatori, che con il loro puntuale lavoro hanno svolto il non facile compito della valutazione delle proposte.

Un ultimo e sentito ringraziamento va all'Università di Udine e a tutti membri del **Comitato di organizzazione locale** che hanno saputo offrire un'accoglienza davvero speciale e una fattiva collaborazione senza la quale non sarebbe stato possibile realizzare questo evento.

Vorrei, inoltre, ringraziare sentitamente il keynote speaker Dino Buzzetti che ci ha onorato con la sua presenza e che tra poco sentiremo nel discorso di apertura; i tre *invited speakers*, che hanno accettato di partecipare nonostante i loro innumerevoli impegni; infine, vorrei ringraziare il numeroso pubblico qui presente (oltre 130 partecipanti) che, ne sono certo, seguirà con molto interesse i lavori

A questo punto, prima di dare avvio ai lavori, non mi resta che augurare a tutti una buona e fruttuosa partecipazione al convegno. Buon lavoro!

Stefano Allegrezza

(dal discorso tenuto in occasione dei saluti iniziali presso il Salone del Parlamento del Castello di Udine, mercoledì 23 gennaio 2019)

# **AIUCD 2019 Panels**

# Models as forms, models as concepts

Øyvind Eide  
Invited speaker

## ABSTRACT

In modelling processes in digital humanities we operationalise our models in the form of computer based artefacts such as programmes, 3D models, or interactive maps. Modelling processes are carried out by one or more persons based on a more or less explicitly defined methodology. This always leave room for making choices on the side of the scholar based on knowledge, abilities, and an understanding of good scholarly practice. In the role of modeller a scholar is a skilled actor doing practical work on mediated objects. The span between the conceptual level and the concrete mediated level will be explored in this paper, based on examples taken from fields of digital humanities practice such as mapping, 3D modelling, and experimental virtual reality systems.



## BIOGRAPHY

Øyvind Eide is a professor in Digital Humanities (Historisch-Kulturwissenschaftliche Informationsverarbeitung) at the University of Cologne. He holds a PhD in Digital Humanities from King's College London (2012) which was funded by a grant from the The Research Council of Norway. His research interests are focused on the modelling of cultural heritage information, especially as a tool for critical engagement with the relationships between texts and maps as media of communication. He is currently engaged in investigating the limitation of texts and maps as means of conveying geographical understanding, using conceptual modelling of texts as his main method. He is also engaged in interdisciplinary theoretical work on modelling as media transformation activity based on a semiotic understanding and was one of the PIs for the project "Modelling Between Digital and Humanities: Thinking in Practice".

# Finding common ground between text, maps, and tables for quantitative and qualitative research

Marieke Van Erp

Invited speaker

## ABSTRACT

Humanities research is increasingly becoming an interdisciplinary affair; digitisation and computing present researchers with interesting opportunities but also with challenges to traditional research methods. Digital archives can be a treasure trove of text, images, maps, tables and more, but accessing these data in an intuitive and interpretable manner takes the humanist into the realm of big data analysis. Whilst daunting at first, this can open up new data perspectives and research collaborations between computer science and the humanities. In this talk, I describe how we forge collaborations between linguists, historians, media studies researchers and computer scientists to answer digital humanities research questions from new, combined perspectives. We use linked data techniques and grounding in time and space to connect data from various archives in various formats and modalities. I will illustrate how this changes the daily practice of researchers, makes us rethink our research methods and I will touch upon some the challenges still ahead.



## BIOGRAPHY

Marieke van Erp uses language technology to make big data more accessible for humanities research. In her research, she has made various datasets more insightful, enabling semi-automatic storytelling from data as well as discoveries of needles in data haystacks, such as shady actors involved in the financial crisis and half-eaten animal specimens. Since 2017, Marieke van Erp is the team leader of the Digital Humanities Lab of the Royal Netherlands Academy of Arts and Sciences Humanities Cluster (KNAW HuC). Van Erp, language technologist and specialist in the field of semantic analysis of texts on entities and events, worked at Tilburg University, Naturalis, and since 2009 at the VU; she pioneered her PhD research (defended in 2010) on digital *humanities methods on historical textual sources*.

# Are the Digital Humanities ‘language insensitive’? Connecting debates about Modern Languages, global cultural representation in DH and the international classroom

Paul Spence  
Invited speaker

## ABSTRACT

Considerable attention has been given to global cultural representation in the Digital Humanities in recent years, but has DH become any more ‘language sensitive’ as a result? Some have proposed that the Modern Languages offer an important new dimension to cultural criticism in DH through what they call a ‘critical DHML’ approach (Pitman & Taylor 2017), but what does such an approach look like in practice? Studying both digital mediations of Modern Languages research and ML perspectives on digital studies, the Language Acts & Worldmaking project’s digital mediations strand explores how a more ML-inflected research agenda might help address cultural diversity in DH, and how this in turn might help to facilitate inclusive pedagogies.



## BIOGRAPHY

Paul Spence is a Senior Lecturer in Digital Humanities at King’s College London and has an educational background in Spanish & Spanish American studies. In the past he has led digital research on a number of projects involving digital edition, user-generated content, innovative visualisation and digital publishing. He co-developed the multi-platform publishing framework xMod (nowKiln), which has been used on over 50 projects. His research currently focuses on digitally mediated knowledge creation, global perspectives on digital scholarship and interactions between modern languages and digital culture. He leads the ‘Digital Mediations’ strand on the AHRC-funded Language Acts and World-making project.



# **AIUCD2019 Panels**

# Metodologie informatiche per le discipline umanistiche: stato dell'arte, criticità e proposte per una classe di laurea magistrale

Francesca Tomasi<sup>1</sup>, Alessandro Lenci<sup>2</sup>, Anna Rovella<sup>3</sup>, Franco Tommasi<sup>4</sup>

<sup>1</sup> Università di Bologna, Italia – francesca.tomasi@unibo.it

<sup>2</sup> Università di Pisa, Italia - alessandro.lenci@unipi.it

<sup>3</sup> Università della Calabria, Italia – anna.rovella@unical.it

<sup>4</sup> Università del Salento, Italia - francesco.tommasi@unisalento.it

## ABSTRACT

Sono solo quattro gli Atenei italiani che hanno ad oggi attiva una laurea magistrale della classe LM43, denominata Metodologie informatiche per le discipline umanistiche: l'Università di Bologna, l'Università della Calabria, l'Università di Pisa e l'Università del Salento. Quattro Atenei che hanno ideato offerte formative differenziate tanto per natura e obiettivi, quanto per caratteristiche e struttura. Le soluzioni proposte meritano dunque un confronto collettivo, che consenta di ragionare sulla forma eterogenea della proposta italiana nel declinare il rapporto fra discipline informatiche e scienze umane. In vista della prossima riforma delle classi dei Corsi di Studio da parte del Consiglio Universitario Nazionale (CUN), pare anche interessante individuare elementi di forza e criticità del decreto ministeriale della LM43.

## PAROLE CHIAVE

ambiti di apprendimento, offerta formativa, figure professionali, classe di laurea.

## INTRODUZIONE

Sono solo quattro gli Atenei italiani che hanno ad oggi attiva una laurea magistrale della classe LM43, Metodologie informatiche per le discipline umanistiche: l'Università di Bologna, l'Università della Calabria, l'Università di Pisa e l'Università del Salento. Si tratta di proposte differenti, elaborate però sulla base di un comune modello, rappresentato dal decreto ministeriale sulle classi di laurea magistrale LM43 [1]. La declaratoria della classe è sufficientemente generica per consentire un ampio margine di interpretazione del ruolo delle metodologie informatiche nelle discipline umanistiche: 48 crediti caratterizzanti, distribuiti fra discipline informatiche (30 cfu), discipline linguistiche e letterarie (10 cfu), discipline storiche, giuridiche, artistiche e archivistiche (8 cfu). E gli obiettivi formativi qualificanti di questa classe vertono sulla creazione di lauree che offrano “le conoscenze metodologiche necessarie per l'uso di strumenti informatici in ambito umanistico” [1]. Metodologia, uso e strumenti: parole chiave che guideranno la nostra riflessione.

Il panel ha lo scopo di condividere le scelte che hanno portato questi quattro Atenei ad attivare una LM di questa classe, ragionando, in particolare, su alcuni aspetti, ovvero cercando di rispondere ai seguenti quesiti:

- come è stata declinata all'atto dell'attivazione la proposta della classe LM43?
- quale rapporto fra i cfu discipline umanistiche (quali ssd) e informatiche (quali ambiti all'interno degli ssd INF/01 e ING-INF/05) + altri ambiti/settori disciplinari (quali)?
- è prevista una modalità speciale di gestione dell'eventuale eterogeneità della provenienza dei laureati triennali?

- qual è il ruolo delle attività complementari (tutorati, tirocini, seminari e laboratori), se contemplate nell'offerta formativa?
- sono state ideate modalità di coinvolgimento delle parti sociali e qual è il loro ruolo nella formazione?
- è stata prevista, o è in previsione, una dimensione internazionale della formazione?
- quali sono gli sbocchi professionali immaginati?

Il panel sarà organizzato per rispondere ai quesiti sopra indicati, descrivendo per ogni realtà nazionale (sezioni 2-5): stato dell'arte della LM attivata, criticità riscontrate nel corso degli anni, proposte future di revisione o integrazione. Un ulteriore giro di tavolo potrà essere necessario per commenti sul decreto della classe LM43 (sezione 6) e per costruire con la comunità di AIUCD un dibattito sulle tematiche affrontate.

### **L'Università di Bologna e la laurea magistrale in Digital Humanities and Digital Knowledge**

L'Università di Bologna ha attivato nell'anno accademico 2017-18 la laurea magistrale internazionale in Digital Humanities and Digital Knowledge (DHDK) [2]. L'idea ispiratrice è stata quella di avviare un percorso formativo internazionale capace di coniugare alcuni saperi umanistici fondativi (filologia, scienze del libro e del documento e linguistica) e saperi tecnologici necessari (pensiero computazionale, tecnologie Web, basi di dati, gestione della conoscenza, multimedia), ibridandoli con il diritto, l'economia e le scienze della comunicazione. L'obiettivo è quello di mettere in condizione il laureato di governare l'intera filiera del lifecycle degli oggetti culturali [3].

La laurea magistrale consente l'accesso, dietro colloquio, a laureati triennali in numerose classi e, fra i 54 cfu obbligatori come requisiti disciplinari, richiede il conseguimento di almeno 18 cfu fra INF/01 o ING-INF/05 e L-FIL-LET/10 o 11. Alla LM43 sono iscritti, al suo secondo anno di vita, circa 60 studenti, metà dei quali internazionali, a fronte di un numero di colloqui di ammissione che si aggira attorno alle 200 unità. La didattica di tutti gli insegnamenti, obbligatori al primo anno e con una serie di possibili opzioni al secondo, è fortemente legata alla realizzazione di progetti pratici, che mettano in condizione lo studente di confrontare il sapere appreso con una possibile spendibilità delle competenze<sup>1</sup>. Ogni anno sono previste attività formative durante il mese di ottobre<sup>2</sup>, anche tenute da professionisti, necessarie a creare la classe, ovvero erogare competenze condivise. La provenienza mista ed eterogenea degli studenti, sia per paese che per formazione, trova nelle Digital Humanities lo strumento utile alla valorizzazione delle conoscenze in ingresso e alla creazione di profili personalizzati.

### **L'Università della Calabria e la laurea magistrale in Gestione e Conservazione dei Documenti Digitali**

L'Università della Calabria ha attivato, nell'anno accademico 2016-2017, la laurea magistrale in Gestione e Conservazione dei Documenti Digitali [4] incardinata nella classe LM43 e proposta dal Dipartimento di Lingue e Scienze dell'Educazione, al cui interno opera dal 1996 il Laboratorio di Documentazione<sup>3</sup>. Il corso di studi è stato istituito in collaborazione con il

<sup>1</sup> DHDK, Published Projects 2017-18, <https://corsi.unibo.it/2cycle/DigitalHumanitiesKnowledge/published-project-2017-18>

<sup>2</sup> DHDK, Attività 2017-18 <https://corsi.unibo.it/2cycle/DigitalHumanitiesKnowledge/notice-board/2017-10-inaugural-opening-meeting>; DHDK, Attività 2018-19 <https://corsi.unibo.it/2cycle/DigitalHumanitiesKnowledge/notice-board/dhdk-introductory-activities-2018-19>

<sup>3</sup> Laboratorio di Documentazione, <http://www.labdoc.it>

Dipartimento di Ingegneria Informatica, Elettronica e Sistemistica con la finalità di costruire un percorso curricolare dedicato alla formazione delle due figure principali della gestione documentale ovvero il Responsabile della Gestione Documentale ed il Responsabile della Conservazione espressamente previste dal Codice dell'Amministrazione Digitale. I due profili professionali sono stati definiti postulando un mix di competenze informatiche, giuridiche e documentali che non trovava riscontro in percorsi universitari all'epoca esistenti. La didattica, che prevede una ampia parte di tirocini ed attività laboratoriali presso i soggetti convenzionati, è centrata su una forte interazione tra saperi umanistici (Archivi digitali, Conservazione dei documenti digitali, Thesauri e tassonomie, Linguistica, ecc), tecnologici (Data e text mining, Basi di dati, Cloud computing) e Giuridico-sociologici (Diritto Amministrativo, Statistica, Sociologia della Pubblica Amministrazione), ed è condotta in stretto collegamento con l'Associazione di Categoria dei Conservatori Accreditati (Assoconservatori) e con AgID che partecipano, entrambi, al comitato di indirizzo. La provenienza da ambiti eterogenei (umanistici ed economico-giuridici) dei 70 studenti iscritti non ha, fino ad ora, creato particolari problemi in ordine alla regolarità dell'apprendimento. L'offerta formativa è completata da un master di secondo livello in Conservatore dei Documenti Digitali che è alla sua quinta edizione. Ad oggi gli studenti che si avviano alla conclusione del percorso di studi hanno, in larga maggioranza, ricevuto offerte di lavoro dai Conservatori accreditati presso i quali avevano effettuato il periodo di tirocinio.

### **L'Università di Pisa e la laurea magistrale in Informatica Umanistica**

Primo corso della classe LM43 in Italia, il corso di laurea in Informatica Umanistica è nato nel 2006 come laurea specialistica, poi trasformata in laurea magistrale nel 2010 [5]. L'offerta formativa è strutturata in 4 percorsi formativi, corrispondenti ad altrettante aree strategiche dell'informatica umanistica, insieme alle rispettive figure professionali:

- Editoria digitale: sono previsti insegnamenti sullo studio e applicazione di strumenti informatici per il campo dell'editoria (gestione e pubblicazione di informazioni strutturate e non strutturate), nonché della biblioteconomia.
- Grafica, interattività, ambienti virtuali: sono previsti insegnamenti atti a fornire competenze che riguardano la comunicazione visuale, l'espressione creativa con l'uso di tecnologie informatiche e la progettazione di soluzioni di interazione sia in ambienti reali che virtuali.
- Management della conoscenza: sono previsti insegnamenti atti a fornire competenze necessarie alla creazione, alla conservazione e alla diffusione di conoscenza all'interno delle organizzazioni pubbliche e private.
- Tecnologie del linguaggio: sono previsti insegnamenti sulle metodologie di produzione di strumenti e risorse per il trattamento automatico della lingua parlata e scritta, strumenti per la ricerca, analisi, estrazione e gestione intelligente dell'informazione contenuta nei documenti testuali, risorse linguistiche.

Tra gli elementi caratterizzanti del corso di laurea si possono citare:

- la stretta continuità con l'omonimo corso triennale, nato nel 2002 e tuttora unico in Italia;
- la sinergia e complementarietà con i corsi di laurea magistrale in Data Science e con Informatica;
- il Seminario di Cultura digitale obbligatorio per tutti gli studenti, nel quale relatori invitati discutono sui contenuti della formazione e della ricerca nel settore dell'Informatica Umanistica;

- i tirocini facoltativi, con più di 100 convenzioni attive, che sono spesso un'occasione per entrare in contatto e stabilire collaborazioni con enti di ricerca e aziende pubbliche e private;
- l'integrazione con l'ecosistema dei centri di ricerca pisani, rappresentato da CNR, Scuola Normale Superiore e Scuola Superiore Sant'Anna, i cui ricercatori contribuiscono all'offerta formativa del corso di laurea.

I valori sul piano occupazionale del corso di laurea sono molto elevati, con un trend temporale generalmente stabile. Dal 2015 al 2017, la percentuale dei laureati che sono occupati in attività regolata da contratto a tre anni dalla laurea non è mai scesa sotto il 67%. Nello stesso arco temporale la percentuale dei laureati occupati a un anno dal conseguimento del titolo è rimasta sempre sopra la soglia del 75%. Questo indica un'ottima capacità dei laureati in Informatica Umanistica di inserirsi rapidamente nel mercato del lavoro.

### **L'Università del Salento e la laurea magistrale in European Heritage, Digital Media and the Information Society**

Il corso di laurea LM43 European Heritage, Digital Media and the Information Society [6] nasce nel 2015 come trasformazione di un precedente master con lo stesso nome condotto congiuntamente alle università di Coimbra (Portogallo), Turku (Finlandia), Graz (Austria) e Colonia (Germania). La collaborazione con le prime tre università partner è tuttora vigente e, secondo la convenzione internazionale, ognuna delle università ha facoltà di inviare i propri iscritti per un semestre presso una delle università partner, con automatico riconoscimento dei crediti ottenuti nel semestre nella sede estera. Stando alla nomenclatura ministeriale, il corso è pertanto definito "a mobilità strutturata". Secondo il progetto, ogni università offre corsi specializzati per un particolare campo di applicazione delle tecnologie digitali ai beni culturali.

L'offerta della nostra università si è mantenuta costante nel passaggio dal master alla laurea magistrale ed è orientata principalmente all'uso delle tecnologie digitali nell'archeologia. L'ammissione al corso richiede agli studenti italiani il possesso di 54 cfu in qualsiasi disciplina umanistica. L'ammissione degli studenti stranieri è invece soggetta alla valutazione di una commissione apposita.

### **La LM43 e il decreto MIUR**

Ci pare utile riportare in chiusura, come strumento utile per la discussione collettiva sulla LM43, quanto riportato dal decreto ministeriale sopra menzionato [1]. I laureati nei corsi di laurea magistrale della classe devono:

- possedere una formazione di base negli studi linguistici, filologici e letterari, o storici e filosofici, o nel campo del patrimonio culturale, o delle arti, musica e spettacolo, o delle scienze della comunicazione, o della formazione;
- essere capaci di affrontare problemi di rappresentazione della conoscenza avvalendosi di appropriati strumenti informatici di analisi;
- possedere gli strumenti teorici, metodologici e tecnici relativi al trattamento informatico dei testi, delle immagini, del suono e del video nell'ambito delle attività di carattere umanistico;
- essere capaci di impostare e realizzare banche dati e archivi digitali con sistemi di gestione negli ambiti specifici di competenza;
- conoscere gli elementi della regolamentazione giuridica sul trattamento degli oggetti digitali e delle tecnologie loro correlate;

- conoscere i principi costitutivi della comunicazione e sicurezza telematica ed essere in grado di utilizzarne pienamente gli strumenti;
- possedere gli strumenti teorici, metodologici e tecnici per il riconoscimento e mantenimento della autenticità e la conservazione degli oggetti digitali;
- essere in grado di utilizzare fluentemente, in forma scritta e orale, almeno una lingua dell'Unione Europea oltre l'italiano, con riferimento anche ai lessici disciplinari.

Vale allora la pena di chiedersi: dati questi obiettivi formativi, che possiamo tradurre in conoscenze e competenze richieste dalla classe, qual è l'approccio disciplinarmente più adeguato alla formazione di professionisti, che siano capaci di cogliere la sfida tanto della ricerca scientifica di settore quanto dell'esigenza di un mercato del lavoro in trasformazione? Detto altrimenti: come tradurre gli obiettivi formativi in settori scientifici, alla luce della naturale evoluzione degli ambiti di ricerca e a fronte dell'esigenza di erogare una didattica consapevole del nuovo ecosistema digitale?

## BIBLIOGRAFIA

- [1] MIUR, Decreti Classi di Laurea Magistrale, [http://attiministeriali.miur.it/media/155598/dmcdl\\_magistrale.pdf](http://attiministeriali.miur.it/media/155598/dmcdl_magistrale.pdf)
- [2] Laurea Magistrale in Digital Humanities and Digital Knowledge (DHDK), <https://corsi.unibo.it/2cycle/DigitalHumanitiesKnowledge>
- [3] Tomasi, F. 2018. *Digital Humanities and Digital Knowledge (DHDK)*. «Umanistica Digitale», 2. DOI= 10.6092/issn.2532-8816/7862.
- [4] Laurea Magistrale in Gestione e Conservazione dei Documenti Digitali, <http://www.labdoc.it/laurea-magistrale-in-gestione-e-conservazione-dei-documenti-digitali/>
- [5] Laurea Magistrale in Informatica Umanistica, <http://www.fileli.unipi.it/infouma/laurea-magistrale/>
- [6] Laurea Magistrale in European Heritage, Digital Media and the Information Society (EUROMACHS), <https://www.letterelinguebbcc.unisalento.it/752;>  
<http://www.euromachs.unisalento.it/index.php/it/>

# Filosofia e Digitale?/Digital Philosophy?

Luca Bandirali, Università del Salento, [luca.bandirali@unisalento.it](mailto:luca.bandirali@unisalento.it)  
Fabio Ciraci,<sup>1</sup> Università del Salento/Centro DH, [fabio.ciraci@unisalento.it](mailto:fabio.ciraci@unisalento.it)  
Luigi Catalani, Università di Salerno, [lcatalani@unisa.it](mailto:lcatalani@unisa.it)  
Francesca Di Donato, Net7, [didonato@netseven.it](mailto:didonato@netseven.it)  
Riccardo Fedriga, [riccardo.fedriga@unibo.it](mailto:riccardo.fedriga@unibo.it)  
Cristina Marras,<sup>2</sup> CNR-ILIESI,<sup>3</sup> [cristina.marras@cnr.it](mailto:cristina.marras@cnr.it)

## ABSTRACT

Il panel si propone, a partire dalle ricerche in corso e da esperienze concrete di dialogo e integrazione tra filosofia e digitale, di ridefinire da un lato, il ruolo importante e specifico che la filosofia ha nel fornire strumenti critici per una comprensione e interpretazione del cambiamento in atto, dei limiti e dei poteri del digitale, e delle prospettive che esso apre sul piano etico, politico, sociale, teorico ed epistemologico. Dall'altro lato, la filosofia dovrebbe contribuire ad aumentare la consapevolezza che non ci si può limitare ad acquisire passivamente strumenti, conoscenze e linguaggi digitali ma occorre avere la capacità sia di usarli consapevolmente, sia di contribuire a crearli attraverso un dialogo interdisciplinare in modo da governare, manipolare, usare la tecnologia nella e per la ricerca. I problemi (e le sfide) nascono, infatti, non solo dalla computazione del dato, ma anche dalla concettualizzazione, rappresentazione e gestione digitale del sapere. La discussione che il panel intende sollecitare è relativa al concreto coinvolgimento delle istituzioni e delle comunità scientifiche a sostegno dello sviluppo sinergico di contenuti e tecnologie, a partire dall'insegnamento e dalla ricerca in filosofia. Solo quando la comunità scientifica sentirà la responsabilità della cura delle risorse online, non le darà per date e solo come funzionali alla ricerca, ma le vedrà, là dove ci sono le condizioni per così dire "ambientali", come una opportunità di collaborazione, di creazione e di riflessione di conoscenza condivisa, solo allora potremmo immaginare di entrare in una dimensione della ricerca aperta, anche in filosofia. Il panel intende aprire un confronto concreto su alcuni specifici aspetti e contesti della ricerca in atto, che fungano da modelli ermeneutici per la ridefinizione dei paradigmi di riferimento e per il rinnovo delle pratiche e delle metodologie della ricerca e delle sue forme narrative e di comunicazione.

## PAROLE CHIAVE

Filosofia Digitale, Ricerca, Infrastrutture, Metodi e Modelli, Scienza aperta.

## INTRODUZIONE

«L'informatique est-elle capable de prendre en charge théorique la totalité du message philosophique?» (Robinet 1969, p. 442). Così si interrogava lo studioso francese André Robinet agli inizi della rivoluzione informatica per le scienze umane quando si nutriva una mitica fiducia nelle "macchine pensanti" e ci si chiedeva appunto se l'informatica fosse capace di prendersi carico della totalità del messaggio filosofico. Con un approccio un po' meno idealista e più pragmatico di fronte alla complessità e alle sfaccettature della "quarta rivoluzione" (Floridi 2014; Roncaglia 2010; Hey *et al.* 2009), possiamo formulare in qualche modo una domanda contraria a quella posta inizialmente: la filosofia è capace di prendersi

---

<sup>1</sup> Coordinatore del panel.

<sup>2</sup> Coordinatrice del panel.

<sup>3</sup> Consiglio Nazionale delle Ricerche - Istituto Lessico Intellettuale Europeo e Storia delle Idee.

carico della totalità del messaggio informatico? Più precisamente qual è il ruolo e il posto che la filosofia occupa nella comunità di pratiche delle *Digital Humanities* (DH)? Come è cambiato il modo di fare ricerca e studiare filosofia con il digitale? Quali sono i progetti in corso? Qual è il contributo del pensiero filosofico alla riflessione sul ruolo assunto dalle nuove tecnologie? E soprattutto, quale ruolo svolto dalla filosofia intesa come teoria critica di fronte ai quesiti posti con urgenza dai new media intesi come oggetti dinamici e trasformativi?

## **FILOSOFIA e DIGITALE? O DIGITAL PHILOSOPHY?**

La filosofia si è aperta, come ha sempre fatto, ai nuovi indirizzi di ricerca, ha assunto nuove definizioni frutto di contaminazioni, di ibridazioni disciplinari e lessicali (si pensi alla “Digital Philosophy” o all’accezione che prendono termini quali “memoria”, “materialità”, “trasparenza”). Data la complessità dei contesti è certamente utile riflettere sui perimetri permeabili che “le filosofie” e l’umanistica digitale tracciano attraverso le loro interazioni, e quali strumenti critici emergono per comprendere l’attuale sistema tecno-sociale. Si tratta di ridefinire le cornici, i parametri e il linguaggio in cui agiamo e pensiamo, anche nella ricerca, alla luce di questa fase e nel contesto di un ecosistema digitale (Marras 2014).

Come le altre discipline umanistiche anche la filosofia ha attraversato l’evoluzione delle tecnologie informatiche per la ricerca, ma gli studiosi sembrano per lo più usare la tecnologia come un mero supporto per la ricerca, al punto che è forse difficile indentificare una comunità di “filosofi digitali” aggregata, che ricorrano in modo coordinato e secondo linee di ricerca e metodologiche condivise allo studio e all’uso delle tecnologie. La tecnologia digitale, in particolare, è usata prevalentemente per leggere e scrivere ed è “relegata” a facilitare l’accesso in archivi (*repository*) e opac, a permettere una più immediata possibilità di pubblicazione e diffusione dei risultati della ricerca. Anche la filosofia, soffre e deve fare i conti, finanche da un punto di vista istituzionale, con un rapporto asimmetrico e una concezione ancillare del digitale, inteso come mero strumento, e non come portatore significativo di un nuovo paradigma culturale (Catalani 2018; Ciraci 2018a; Marras 2017).

Riteniamo importante affermare che la relazione è invece complessa e genera un mutuo cambiamento nelle metodologie e nelle interpretazioni. La questione, infatti, è come concettualizzare questo processo di interscambio e come tematizzare l’euristica contenuta nella relazione tra filosofia e pratiche computazionali.

## **OBIETTIVI**

Il panel si propone, a partire dalle ricerche in corso e da esperienze concrete di dialogo e integrazione tra filosofia e digitale, di ridefinire da un lato, il ruolo importante e specifico che la filosofia ha nel fornire strumenti critici per una comprensione e interpretazione del cambiamento in atto, dei limiti e dei poteri del digitale e delle prospettive che esso apre sul piano etico, politico, sociale, teorico ed epistemologico. Dall’altro lato, la filosofia dovrebbe contribuire a formare o aumentare la consapevolezza che non ci si può limitare ad acquisire passivamente strumenti, conoscenze e linguaggi digitali ma occorre avere la capacità sia di usarli consapevolmente, sia di contribuire a crearli attraverso un dialogo interdisciplinare in modo da governare, manipolare, usare la tecnologia nella e per la ricerca. I problemi (e le sfide) nascono, infatti, non solo dalla computazione del dato, ma anche dalla concettualizzazione, rappresentazione e gestione digitale del sapere.



## PROSPETTIVE

La discussione che il panel intende sollecitare è relativa al concreto coinvolgimento delle istituzioni e delle comunità scientifiche a sostegno dello sviluppo sinergico di contenuti e tecnologie, a partire dall'insegnamento e dalla ricerca in filosofia. Solo quando la comunità scientifica sentirà la responsabilità della cura delle risorse online, non le darà per date e solo come funzionali alla ricerca, ma le vedrà, là dove ci sono le condizioni per così dire "ambientali", come una opportunità di collaborazione, di creazione e di riflessione di conoscenza condivisa, solo allora potremmo immaginare di entrare in una dimensione della ricerca aperta, anche in filosofia (cfr. Spampinato, *et al.* 2018, Di Donato 2009). Così come per tutte le altre discipline umanistiche il lavoro da fare è tanto. Si tratta di individuare percorsi d'azione e di intervento, con un percorso specifico, e soprattutto attivarsi per superare le barriere per una effettiva ed efficace dialettica tra ricerca filosofica e digitale, e attivare e tematizzare sempre più frequenti zone di intersezione e osmosi, progetti, ricerche, pubblicazioni, piattaforme interoperabili. Un lavoro questo che non può essere affidato ai singoli ma alla cooperazione tra studiosi e in cui la formazione gioca un ruolo centrale. Pertanto, il panel intende aprire un confronto concreto su alcuni specifici aspetti e contesti della ricerca e dell'insegnamento della "filosofia digitale" e della "filosofia per il digitale" in atto, che fungano da modelli ermeneutici per la ridefinizione dei paradigmi di riferimento e per il rinnovo delle pratiche e delle metodologie della ricerca e delle sue forme narrative e di comunicazione.

### ABSTRACT (in ordine alfabetico)

**Luca Bandirali**, Nuovi abiti per la vecchia cerimonia: Aristotele, Virgilio, Shakespeare e la narrativa digitale.

In una visione storicistica delle forme narrative, non c'è alcuna relazione fra la tragedia attica e una serie televisiva. Al contrario, in una visione metafisica, l'Odissea e Apocalypse Now raccontano la stessa storia; d'altra fu Umberto Eco a sostenere che il meccanismo di Ombre rosse era stato già descritto da Aristotele. Se le strutture profonde del racconto cinematografico corrispondono effettivamente per buona parte a quelle individuate nella Poetica, che ne è di Aristotele nell'epoca della narrativa digitale? Virgilio ha mai scritto una puntata di Gomorra? E cosa ci fa Amleto sull'holodeck? Le storie sono sempre le stesse o i nuovi media (e i nuovi fruitori) le hanno modificate per sempre?

**Luigi Catalani**, *Dalla filosofia all'informatica, e ritorno: dialogo ermeneutico e didattica wiki.*

La tecnologia wiki può essere l'ambiente digitale in cui sperimentare nuove modalità di trasmissione e produzione di contenuti, lo spazio interattivo dove costruire, discutere e comunicare contenuti e metodi filosofici delle *Digital Humanities*. L'intervento prende spunto dall'esperienza didattica pluriennale del corso di *Informatica applicata alle scienze filosofiche*, condotta presso l'Università di Salerno, in particolare dagli esiti di una sperimentazione che ha coinvolto circa 120 studenti nella redazione di un manuale disciplinare che muove dalle radici filosofiche dell'informatica per poi percorrere il percorso inverso, soffermandosi sulle nuove questioni filosofiche innescate dalla diffusione pervasiva delle ICT e di Internet. Un terreno epistemologico comune all'informatica e alla filosofia può individuarsi non solo nella filosofia dell'informazione, ma anche in quegli aspetti della *Library and Information Science* che si concentrano sui temi dell'organizzazione della conoscenza e della categorizzazione del reale. Le piattaforme wiki si scoprono funzionali a una riflessione sul rapporto tra la filosofia e l'informatica che passa per uno sforzo condiviso e aperto di concettualizzazione, costruzione e negoziazione dei contenuti (*Open Science*). Lo

sviluppo di competenze informative digitali permette agli studenti di misurarsi con le peculiarità speculative, le implicazioni etiche e sociali della cultura digitale, ovvero di esercitare il pensiero critico elaborando nuove interpretazioni della complessità.

**Fabio Ciraci**, *Il diavoletto di Floridi. Il problema del verosimile.*

La discussione avviata da Luciano Floridi sulla *infosfera* pone dei quesiti urgenti circa il ruolo che le discipline umanistiche sono chiamate ad assolvere nei confronti dell'inquinamento del sistema informativo, con rilevanti implicazioni in ambito etico e sociale, in relazione all'autorevolezza delle fonti scientifiche in rete. Particolare rilievo assume la questione della *veridicità* nell'ambito dell'informazione semantica fattuale: a partire dall'impianto presentato ne *La quarta rivoluzione*, si intende allora discutere del ruolo svolto dalla *logica della verosimiglianza*, la quale sembra sfuggire allo schema proposto da Floridi, risultando impossibile ridurre il verosimile alla cattiva informazione (informazione non veritiera e non intenzionale) o alla disinformazione (informazione non veritiera ma intenzionale). Si intende allora verificare in che modo e con quali limiti il paradigma della post-verità suggerito da Maurizio Ferraris permetta o meno di indagare, con maggiore profitto, la logica della verosimiglianza, per comprendere la natura dei processi che presiedono, per esempio, alla diffusione delle *fake news*. Si cercherà infine di mostrare come “il problema del verosimile” richieda un nuovo approccio teoretico, una nuova teoria che tenga conto del cambiamento non solo dell'oggetto (nel senso del contenuto semantico informativo di una informazione) ma anche delle condizioni di possibilità ricettive del soggetto conoscitivo e della comunicazione in ambiente digitale.

**Francesca Di Donato**, *Filosofia digitale e uso pubblico della ragione: il paradigma della scienza aperta.*

Il paradigma della scienza aperta, che si è affermato negli ultimi anni, sta diventando uno standard per la comunicazione scientifica a livello europeo e internazionale, come riconosciuto dalle policy della Commissione Europea e dal Plan pubblicato nell'ottobre di quest'anno. Esso implica una trasformazione metodologica nel modo di fare ricerca, trasformazione che si fonda sull'uso delle tecnologie digitali, che sono considerate non solo come un mezzo al servizio della ricerca vera e propria, ma piuttosto come parte essenziale del modo di fare scienza. Il presente contributo intende mettere in luce come tale impostazione abbia importanti basi filosofiche e soffermarvisi, a partire dalle seguenti domande.

Quali sono i concetti e le teorie filosofiche a fondamento dell'open science? Come tali concetti introducono un radicale cambio di paradigma nella tradizionale metodologia di ricerca nelle scienze umane? E in che modo la filosofia è in grado di acquisire un ruolo determinante nell'informare i processi della ricerca, sia nelle scienze umane e sociali e sia nelle cosiddette scienze “dure”? Infine, ci si concentrerà sull'analisi di *workflow* open science specificatamente nell'ambito dell'umanistica digitale, per proporre esempi concreti di tale approccio.

**Riccardo Fedriga**, *Autori, testi controfattuali, esistenza di mondi, libertarismo e funzioni di status nell'ecosistema della narrazione digitale.*

Partendo dalla nozione di funzione di status, il mio intervento intende esplorare la sua applicazione all'infosfera digitale, in particolare per quanto concerne la relazione con l'autorialità; ii) la creazione di storie che si caratterizzano come controfattuali operati dal lettore/utente rispetto a un testo dato e/o un oggetto di valore (i.e. sociale) dato; iii) la possibilità, in particolare, da parte di autori e/o di lettori di fare diversamente rispetto a tali iscrizioni attraverso la creazione di mondi narrativi (*worldbuilding*). John Searle spiega esaurientemente la definizione di “funzione di status” assumendo che lo “status” è fonte di

potere, e nel caso dell'utenza online tale potere è essenziale che si manifesti: per Searle gli esseri umani possiedono la capacità di creare una realtà fatta di relazioni di potere “che esiste soltanto perché questi stessi esseri umani hanno creduto che quella realtà esistesse; in una simile realtà le relazioni di potere emergono dal fatto che a una certa persona o a un certo oggetto è stato assegnato un certo status e che attraverso quello status gli è stata assegnata una funzione che può essere performata solo in virtù dell'accettazione collettiva di quello status. Una volta data per innata la tendenza ad attribuire del valore in un gioco volto a costruire il proprio potere, si veda in campo una forza costituita dall'autore, dall'oggetto di valore che è il suo contenuto online e da chi amministra tale mercato, l'utenza. Ne deriva che l'esistenza – e la sussistenza – in rete viene vista come una sorta di dialettica tra i criteri di esistenza di un dato testo sul web (un e-book, un post, un prodotto qualsiasi), la sua elaborazione cognitiva da parte dell'utenza (personalissima e, va ripetuto, potente) e la tecnologia che la veicola, sancendo ancora la distinzione tra il contenuto e il suo supporto, ma ora in aperto confronto con ciò che genera potere, “funzioni di status”: il bacino di utenti/lettori. Le funzioni di status appaiono dunque determinanti nella negoziazione della libertà rispetto al determinismo digitale: per questo è di capitale importanza indagarne le applicazioni.

**Cristina Marras**, *Filosofia digitale: un approccio plurilinguistico – introduzione al panel*.

La filosofia è ricerca specialistica, storica e teoretica, che guarda alla tradizione e allo sviluppo del pensiero attraverso i testi, i contenuti, il linguaggio, e che non può trascurare e deve assumersi la responsabilità di interpretare la complessità della nostra contemporaneità. A partire dall'esperienza maturata nell'ambito dell'interoperabilità delle infrastrutture digitali per la ricerca in filosofia e della modellizzazione della conoscenza a partire dalla ricerca terminologica e lessicale nella storia delle idee, nell'intervento mi propongo di riflettere sulla necessità, sempre più urgente, che la filosofia contribuisca alle discussioni che caratterizzano le DH e che coinvolgono le strategie per una accessibilità intelligente alla conoscenza e ai dati, le modalità di rappresentazione della conoscenza, la produzione di nuova conoscenza integrata e non modulare. È nel governo dei processi che la filosofia offre strumenti utili ed efficaci, ma questa capacità risiede nel rapporto che si riesce a stabilire tra conoscenza e competenza e nella possibilità intrinseca al pensiero filosofico di tradurre le domande in metodologie, categorie, modelli epistemologici per orientarsi, per capire i problemi e per governare i processi.

## **BREVE BIOGRAFIA DEI PARTECIPANTI AL PANEL (IN ORDINE ALFABETICO)**

Luca Bandirali, Università del Salento, insegna Teorie e tecniche del linguaggio audiovisivo presso il corso di laurea DAMS e presso il corso di laurea in Scienze della Comunicazione dell'Università del Salento; fa parte del comitato di redazione della rivista di critica cinematografica “Segnocinema”; è membro del Centro Studi e Ricerca nel settore del cinema e dell'audiovisivo della Fondazione Apulia Film Commission; è membro della Commissione Tecnica di Valutazione dell'Apulia Film Fund; fa parte del Gruppo di Ricerca Interdipartimentale in Digital Humanities dell'Università del Salento. Web: <https://www.unisalento.it/scheda-utente/-/people/luca.bandirali> - email: [luca.bandirali@unisalento.it](mailto:luca.bandirali@unisalento.it)

Luigi Catalani, è bibliotecario professionista presso la Biblioteca Provinciale di Potenza; responsabile progetti educativi Polo delle Arti e della Cultura della Provincia di Potenza; docente di *Informatica applicata alle scienze filosofiche* presso l'Università di Salerno; responsabile sezione *Digital Humanities* del Centro FiTMU dell'Università di Salerno; coordinatore nazionale “Wikimedia Italia scuola e università”; membro del Gruppo di studio

AIB sull'information literacy. Web: [https://it.wikipedia.org/wiki/Utente:Luigi\\_Catalani](https://it.wikipedia.org/wiki/Utente:Luigi_Catalani) - email: [lcatalani@unisa.it](mailto:lcatalani@unisa.it)

Fabio Ciraci (proponente e coordinatore del Panel), lavora presso il Dipartimento di Studi Umanistici dell'Università del Salento, è Professore Aggregato di *Storia della Filosofia Italiana e di Informatica Umanistica*, Facoltà di Lettere e Filosofia, Lingue e Beni Culturali; direttore scientifico del gruppo di ricerca Digital Humanities dell'Unisalento; segretario del *Centro interdipartimentale di ricerca su Arthur Schopenhauer e la sua scuola*; executive Editor della Rivista *H-ermes. Journal of Communication*; dal 2018 è membro del Direttivo dell'AIUCD. Web: <https://www.unisalento.it/scheda-utente/-/people/fabio.ciraci> - email: [fabio.ciraci@unisalento.it](mailto:fabio.ciraci@unisalento.it)

Francesca Di Donato, lavora per la SME Net7 srl, di Pisa, dove coordina il dipartimento dedicato agli strumenti innovativi e ai servizi per la comunicazione della ricerca, in particolare le *Digital Humanities* e il patrimonio culturale. È stata ricercatrice presso l'Università di Pisa e la Scuola Normale Superiore; è abilitata come Professore Associato per l'insegnamento di Filosofia Politica; dal 2017 è membro del direttivo di AISA - *Associazione Italiana Scienza Aperta*; è membro di IOSSG - *Italian Open Science Support Group*; dal 2015 al 2018 ha fatto parte di *Europeana Association Network Councillor*. Web: <http://orcid.org/0000-0003-0144-8934> - email: [didonato@netseven.it](mailto:didonato@netseven.it)

Riccardo Fedriga, Professore associato, Dipartimento di Filosofia e Comunicazione dell'Università di Bologna; membro del comitato di direzione della Rivista "Dianoia"; è curatore responsabile di numerosi e importanti progetti editoriali per le modalità di lettura ibride dei nativi digitali; direzione scientifica della Collana: *Performing Ideas*, Encyclomedia Publishers sulla storia delle idee e della lettura; è membro della *Società italiana per lo studio del Pensiero medievale* (SISPM) e della *Società Internazionale per lo Studio della Filosofia Medievale* (SIEPM). Web: <https://www.unibo.it/sitoweb/riccardo.fedriga/> - email: [riccardo.fedriga@unibo.it](mailto:riccardo.fedriga@unibo.it)

Cristina Marras (proponente e coordinatrice del Panel), è ricercatrice presso l'Istituto Lessico Intellettuale Europeo e Storia delle Idee del Consiglio Nazionale delle Ricerche di Roma; co-coordinatrice del progetto *Modelling between Digital and Humanities: Thinking in Practice*; responsabile scientifico per l'ILIESI del progetto PON "Beni e Patrimonio Culturale: governare la progettazione"; dal 2010 al 2011 è Science Officer presso l'European Science Foundation di Strasburgo per il peer review, la valutazione e la qualità della ricerca; dall'anno accademico 2018/19 tiene il corso in convenzione *Digital Humanities for Philosophy* al dipartimento di Filosofia della Sapienza Università di Roma, nel quale, sempre in convenzione, dal 2007 ha insegnato "Teoria della Comunicazione" e "Semiotica"; è membro della redazione della rivista *Umanistica Digitale* e di *Lexicon Philosophicum*; è stata membro del direttivo dell'AIUCD dal 2014 al 2017, e della "Sodalitas Leibnitiana dal 2010-2016. Web: <http://www.iliesi.cnr.it/profilo.php?name=as> - email: [cristina.marras@cnr.it](mailto:cristina.marras@cnr.it)

## BIBLIOGRAFIA

- [1] Bandirali, L. 2016. Origine, connessione, trasformazione nelle strutture narrative del cinema e della serialità televisiva. In: *Fata morgana*, anno X n. 29, pp. 155-169.
- [2] Bandirali, L., Terrone E. 2009. Il sistema sceneggiatura. Scrivere e descrivere i film. Torino, Lindau.

- [3] Bandirali, L., Terrone E. 2018. *Tele-machia*. In: M. Guerra, S. Martin, S. Rimini (a cura di), *Universo Gomorra. Da libro a film, da film a serie*, Milano - Udine, Mimesis, pp. 19-26.
- [4] Bynum, T. W., Moor, J. H. (eds) 1998. *The Digital Phoenix how computers are changing philosophy*, Oxford University Press, Oxford.
- [5] Castellucci, P. 2009. Dall'ipertesto al web. Storia culturale dell'informatica, Laterza, Roma-Bari.
- [6] Catalani, L. 2018. Informatica umanistica e Digital Humanities. In: *Medium*, 19/2/2018, <https://medium.com/@luigicatalani/informatica-umanistica-e-digital-humanities-ff57c44d68be>
- [7] Catalani, L. 2018. I progetti Wikimedia per l'educazione civica digitale – le biblioteche digitali nella didattica. In: *Bricks*, Agosto 2018, pp. 62-66, [http://www.rivistabricks.it/wp-content/uploads/2018/08/2018\\_3\\_10\\_Catalani.pdf](http://www.rivistabricks.it/wp-content/uploads/2018/08/2018_3_10_Catalani.pdf)
- [8] Catalani, L. 2018b. Informatica e filosofia: contatto!. In: *Medium*, 23/2/2018, <https://medium.com/@luigicatalani/informatica-e-filosofia-contatto-5dcc2d813cc2>
- [9] Catalani, L. 2018c. Le origini filosofiche del paradigma computazionale. In: *Medium*, 2/3/2018, <https://medium.com/@luigicatalani/le-origini-filosofiche-del-paradigma-computazionale-e885c8c5a65f>
- [10] Catalani, L. 2017. I progetti Wikimedia per l'apprendimento delle competenze informative e digitali in biblioteca, a scuola, nelle università. In: *AIB Studi*, 2 (2017), pp. 253–263, <http://aibstudi.aib.it/article/view/11654>
- [11] Catalani, L. 2016. «Omnia numerorum videntur ratione formata». A 'Computable World' Theory in the Early Medieval Philosophy. In: *History and Philosophy of Computing. Third International Conference, HaPoC 2015, Pisa, Italy, October 8–11, 2015. Revised Selected Papers*, ed. by F. Gadducci — M. Tivosanis, Springer, Pisa 2016 (IFIP Advances in Information and Communication Technology, 487), pp. 131–140.
- [12] Ciraci, F. 2018. *Digital Humanities: fra Kant e McLuhan (e strumenti trasformativi)*, <http://cotidieblogger.blogspot.it/2018/02/digital-humanities-fra-kant-e-mcluhan-e.html>, 24.02.2018.
- [13] Ciraci, F. 2018a, Presentazione al Panel: Common Strategies and Aims of Italian Reserach Centers for Digital Humanities, 7th AIUCD Annual Conference, Patrimoni Culturali nell'era digitale. Memorie Culture Umanistiche e Tecnologiche, Bari 31 gennaio - 3 febbraio 2018, *Book of Abstracts*, a cura di D. Spampinato, et al. <http://www.aiucd2018.uniba.it>
- [14] Ciraci, F. 2018b. Feticci digitali e memorie parallele. Una proposta minimalista. In: *Le risorse digitali per la storia dell'arte moderna in Italia*, a cura di F. Conte, Edizioni Storia e Letteratura, Roma, pp. 149-158.
- [15] Davis, M. 2000. *The Universal Computer. The Road from Leibniz to Turing*, W. W. Norton & Company, New York-London.
- [16] Di Donato, F. 2009. La scienza e la rete. L'uso pubblico della ragione nell'età del web, FUP, Firenze.
- [17] Di Giandomenico, M. (a cura di). 2000. *L'uomo e la macchina trent'anni dopo. Filosofia e informatica ieri ed oggi*. Atti del Convegno nazionale della Società Filosofica Italiana, Bari, 24-26 ottobre 1997, Bari.
- [18] Floridi, L. 2014. *The Fourth Revolution. How the Infosphere is Reshaping Human Reality*, Oxford University Press, Oxford.
- [19] Floridi, L. (ed.) 2014. *The Blackwell Guide to Philosophy of Computing and Information*, Oxford University Press, Oxford.
- [20] Hey, T., Tansley S., Tolle K. (eds.) 2009. *The fourth Paradigm. Data-Intensive Scientific Discovery*, Microsoft research. Redmon, Washington.

- [21] Marras, C. 2018, Presentazione al Panel: Common Strategies and Aims of Italian Research Centers for Digital Humanities, 7th AIUCD Annual Conference, Patrimoni Culturali nell'era digitale. Memorie Culture Umanistiche e Tecnologiche, Bari 31 gennaio – 3 febbraio 2018, *Book of Abstracts*, a cura di D. Spampinato, et al.  
<http://www.aiucd2018.uniba.it>
- [22] Marras, C. 2018a. Filosofia, ricerca e innovazione. Considerazioni a margine del documento “Orientamenti per l'apprendimento della filosofia nella società della conoscenza”. In: *Comunicazione Filosofica*, 40, maggio 2018, p. 9,  
<https://www.sfi.it/files/download/Comunicazione%20Filosofica/cf40.pdf>
- [23] Marras, C. 2017, DH and Philosophy: digital ancilla philosophiae? Or the philosophy of the snail, intervento all'Workshop – Methodology in the Digital Humanities, Master in Digital Humanities and Knowledge, Bologna, 10 novembre 2017
- [24] Marras, C. 2014. Exploring digital environments for research in philosophy. Metaphorical models and sustainability, In: *AIUCD '14, Proceedings of the Third AIUCD Annual Conference on Humanities and Their Methods in the Digital Ecosystem*, Bologna, September 18-19 2014, edited by F. Tomasi, R. Rosselli del Turco, A.M., Tammaro, ACM, New York, doi:10.1145/2802612.2802639
- [25] Pagallo, U. 2005. Introduzione alla filosofia digitale. Da Leibniz a Chaitin, Giappichelli, Torino.
- [26] Robinet, A. 1969. La communication philosophique à l'ère des ordinateurs. In: *Revue Internationale de Philosophie*, XXIII, 1969.
- [27] Roncaglia, G. 2010. La quarta rivoluzione. Sei lezioni sul futuro del libro, Roma – Bari, Laterza.
- [28] Spampinato, D. et al. (a cura di). 2018. Book of Abstracts, 7th AIUCD Annual Conference, Patrimoni Culturali nell'era digitale. Memorie Culture Umanistiche e Tecnologiche. Bari 31 gennaio – 3 febbraio 2018,  
<https://umanisticadigitale.unibo.it/pages/view/boa>

Ultimo accesso agli URL: 4 gennaio 2019

# L'uso degli applicativi HGIS-WebGIS nella ricerca geo-storica

Tiago Luis Gil<sup>1</sup>, Massimiliano Grava<sup>2</sup>, Nicola Gabelliere<sup>3</sup>, Vinicius Maluly<sup>4</sup>

<sup>1</sup> Università di Brasilia / Università Ca'Foscari di Venezia (Visiting Scholar), Italy, tiagoluisgil@gmail.com

<sup>2</sup> Università di Pisa, massimiliano.grava@cfs.unipi.it

<sup>3</sup> Università di Trento, n.gabellieri@hotmail.com

<sup>4</sup> Ecole des Hautes Etudes en Sciences Sociales, vmaluly@gmail.com

## ABSTRACT

Il presente panel mira a presentare e discutere modelli di GIS spaziotemporali, mettendo in comunicazione ricercatori provenienti da settori diversi, senza confini cronologici o disciplinari. In un'ottica di *public history*, si vogliono infatti condividere le migliori pratiche e le strategie più innovative – ma anche le problematiche – sull'uso dei sistemi informativi territoriali. La cartografia digitale è universalmente riconosciuta come strumento euristico per lo studio del passato, sia nella vettorializzazione della cartografia storica, sia per la georeferenziazione di dati provenienti da altre fonti; allo stesso tempo, gli strumenti GIS si sono dimostrati un potente mezzo di applicazione e condivisione dei dati, sia per la pianificazione territoriale, sia per la comunicazione a una grande platea. La rappresentazione cartografica virtuale del passato si presenta quindi come uno dei nuovi punti di incontro tra ricerca e disseminazione.

## PAROLE CHIAVE

Historical-GIS, WebGIS, Cartography, History

“Geography is the study of spatial differentiation, history the study of temporal differentiation. Historical GIS provides the tools to combine them to study patterns of change over space and time” (Knowles, 2002).

Nel 2002 Knowles, insieme a un nutrito gruppo di studiosi tutti ascrivibili in vario modo a settori disciplinari umanistici, formulò quella che forse sarebbe stata la più efficace definizione di GIS storico. Un assioma divenuto, negli anni successivi, principio di base per quanti si sarebbero successivamente dedicati a sudi oggi catalogati HGIS. Non solo quindi strumenti di produzione cartografica quanto piuttosto dei database geografici.

È oramai un dato di fatto che i sistemi informativi geografici, pur con qualche affanno, si stanno affermando anche nel settore della ricerca storica; specie in quelle discipline a carattere quantitativo direttamente connesse con lo studio del territorio.

La filiera storica, che oggi come in passato inizia tra la polvere degli archivi e con l'ausilio di fonti bibliografiche per poi proseguire con lo studio dell'informazione prodotta, si conclude però, nel caso di ricerche di storia digitale, con la pubblicazione sul Web di una serie di livelli informativi tematici. Un costrutto storico che è edito su piattaforme in cui, questa la grande differenza rispetto a ricerche tradizionali, posso coesistere le fonti originali, le metafonti e le interpretazioni dello storico. La costruzione di queste banche dati geografiche non deve quindi essere finalizzata alla realizzazione di mappe statiche «a-spaziali» quanto piuttosto all'elaborazione di una cartografia numerica che, in quanto digitale, è in grado di assumere dinamicità.

La cartografia digitale è per questo motivo universalmente riconosciuta come strumento euristico per lo studio del passato, sia nella vettorializzazione della cartografia storica, sia per

la georeferenziazione di dati provenienti da altre fonti; allo stesso tempo, gli strumenti GIS si sono dimostrati un potente mezzo di applicazione e condivisione dei dati, sia per la pianificazione territoriale, sia per la comunicazione a una grande platea. La rappresentazione cartografica virtuale del passato si presenta quindi come uno dei nuovi punti di incontro tra ricerca e disseminazione.

Se dal lato l'uso dello strumento e il ragionamento sulle sue problematiche si sta sviluppando, dal lato pubblico si stanno creando strumenti sempre più amichevoli per consentire ai meno esperti di aggiungere dati, creare proprie mappe, condividere database geostorici. Il punto di incontro di questi due mondi non è tuttavia sempre agevole.

Posizionare un dato storico all'interno di uno spazio fisico consente quindi di comprendere i fatti studiati al di là di ciò che può essere ottenuto semplicemente studiando il fenomeno investigato senza il contesto spaziale in cui si trova e in cui questo si trova e si è sviluppato. I software GIS, con cui viene prodotta la cartografia numerica, consentono non solo di svolgere questo compito. Strumenti quantitativi consentendo che consentono di ubicare nello spazio dei layer informativi e, se si parla del lavoro degli storici, di poter interrogare la serie informative con i procedimenti geostatistici.

Il concetto di metafonti, introdotto per la prima volta da Jean-Philippe Genet nel 1992, fa riferimento proprio alla produzione di nuove fonti "immateriali". Questi, prodotti grazie al supporto informatico, non sono semplicemente trasposizioni di documenti da "analogico" a "digitale", quanto piuttosto il risultato della costruzione "ex novo" di nuovi documenti.

Il presente panel mira quindi a presentare e discutere modelli di GIS spaziotemporali, mettendo in comunicazione ricercatori provenienti da settori diversi, senza confini cronologici o disciplinari. In un'ottica di *public history*, si vogliono infatti condividere le migliori pratiche e le strategie più innovative – ma anche le problematiche – sull'uso dei sistemi informativi territoriali; ma anche una approfondita discussione sul il concetto di metafonte e sull'uso di sistemi di informazione geografica nella storia. Nello specifico quindi cercheremo di evidenziare come l'aumento del numero di indagini la cui metodologia sia associata al geoprocessing e ai sistemi di informazione geografica consenta altresì una costante rilettura e correzione delle procedure e degli standard connesse agli Historical-GIS.

Un altro importante dibattito riguarda la creazione di mappe narrative e animazioni come possibili forme di rappresentazione della conoscenza storica. È possibile rappresentare la complessità della storia attraverso l'uso di mappe? Le mappe dinamiche e animate sarebbero modi privilegiati per presentare i risultati della ricerca? In che misura gli strumenti della geostatistica sarebbero efficaci nel comunicare i risultati della ricerca nella storia?

La proposta di questo panel prende pertanto spunto da queste premesse ponendo particolare attenzione ed enfasi sulle seguenti tematiche: Possibilità e limiti d'impiego dei GIS nella ricerca storica; Soluzioni ai problemi di acquisizione e gestione di informazioni descrittive e di quelle cartografiche; La pubblicazione dei geodati su piattaforme: Cloud-GIS e WebGIS; L'uso di HistoricalGIS per la costruzione di piattaforme collaborative; GIS e geografia storica tra ricerca e applicazioni pratiche. Tutti questi sono problemi che devono essere affrontati dall'insieme di relatori scelti per il panel, che guarda i diversi volti in questo nuovo campo di conoscenza. I periodi coperti sono diversi e questo contribuisce alla valutazione dei diversi significati dell'uso dei Sistemi di Informazione Geografica nella ricerca in storia.

## **PARTECIPANTI:**

### **Massimiliano Grava, Le strade della Piana di Pisa in epoca leopoldina. Una ricerca di geografia storia realizzata con applicativi GIS**

Questa ricerca di geografia storica, basata sull'esame del caso studio della viabilità pubblica del territorio comunitativo pisano alla data del 1847, ci restituisce in realtà un significativo



spaccato di quel quadro geopolitico toscano relativo alle strutture amministrative del restaurato Granducato intorno alla metà del XIX secolo. Una stagione, quella Toscana durante il governo di Leopoldo II d'Asburgo-Lorena (1824-1859), che si caratterizza per l'introduzione di una straordinaria serie di riforme della macchina burocratica granducale. Riordini normativi di una portata tale che l'intero apparato Statale si riorganizza tanto a livello centrale che in quello periferico in forme maggiormente efficienti. La trasposizione dei dati archivistico-cartografici in ambiente GIS ha infine consentito di ricostruire, anche cartograficamente, i cambiamenti della viabilità di questo territorio. Un costrutto digitale edito su una piattaforma WebGIS su cui coesistono le fonti originali, le metafonti elaborate e la cartografia numerica digitale prodotta dagli enti pubblici. Obiettivo di questo saggio è pertanto lo studio della viabilità del pisano in epoca lorenesse quale elemento conoscitivo del paesaggio – si pensi ad esempio alla questione dell'onomastica storica o di tutte le strade dismesse o modificate tra la metà del XIX secolo ed oggi – ma anche quello di dotare di un efficace strumento coloro che si occupano di pianificazione territoriale riguardante il sistema viario, siano questi amministratori della cosa pubblica, tecnici o privati cittadini.

### **Tiago Gil, Narrative Maps in History**

Maps have long been used to represent historical knowledge. Lately, this use has been widely diffused, with the publishing of several new works which emphasize cartography as a language for historians. The approaches are diverse and the research problems addressed to these methodologies are numerous. The purpose of this text is to highlight a specific type of cartographic production in history: the use of narrative maps. Under the expression "narrative maps" there are at least two possibilities of approach. The first concerns the mapping of narratives, that is, converting texts to maps, something close to Franco Moretti's work called "Atlas of the European Novel". In this case, the texts are read and, through various procedures, decomposed in the form of maps, from certain variables chosen by the one who analyzes. The second approach concerns the representation of the results of the historians' researches in the form of maps oriented simultaneously in time and space. Although this visualization is commonly associated with economic data, it can also be used to describe and analyze other processes. There are at least six two-dimensional forms of cartographic representation of narratives: a) series of images (like cinema), b) sequential arts (like comics), c) vector animations, d) small multiples, e) anamorphosis, f) traditional static maps, provided they have features that represent the narrative, usually through arrows, lines or other elements that indicate spatio-temporal dynamics. Using authors like Jacques Bertin and Edward Tufte, we pretend to debate the usage of narrative maps in historical research.

### **Nicola Gabellieri, Dalle foto al GIS. Il software WSL Monoplotting Tool per la georeferenziazione e la vettorializzazione di foto oblique da terra per una geografia storica applicata**

Secondo Tosh (2008), la Public History è quel settore della ricerca dedicato a condurre di ricerche di interesse sia per gli storici accademici sia per il grande pubblico. In questo contesto, la Storia Applicata fornisce dati storici a istituzioni e altre discipline come, tra le altre, la pianificazione e la preservazione e valorizzazione del patrimonio e del paesaggio. Nel campo della geografia storica, la ricostruzione di paesaggi rurali del passato è stata spesso basata sull'analisi di cartografie e fotoaeree, così come di foto prese dal terreno. Mentre l'uso di fotoaeree e cartografie hanno beneficiato dalla diffusione dei software GIS, l'utilizzo delle fotografie è stato limitato a una valutazione qualitativa dei cambiamenti del paesaggio a grande scala. Questa presentazione mostra il potenziale di un software sviluppato dal gruppo di ricerca del Institute for Snow and Avalanche Research (WSL) svizzero per georeferenziare

e vettorializzare fotografie storiche. L'uso del software Monoplotting Tool apre la possibilità di nuove applicazioni per l'uso di foto topografiche e il loro uso per misurare i cambiamenti nella copertura vegetale. La presentazione illustra il nuovo software e esamina come può essere usato per esaminare paesaggi montani in Liguria e in Trentino nel XIX e XX secolo. L'uso di fotografie storiche di terreno come fonti per lo studio dei cambiamenti sull'uso e copertura del suolo può offrire numerosi vantaggi, e arricchire le analisi storiche basate sul GIS per la gestione e la pianificazione del paesaggio culturale.

### **Vinicius Maluly, An analysis on textual sources for cartographic and territorial understanding: the case of Auguste de Saint-Hilaire (1816-1822)**

The connection between cartographical productions and territorial understandings is close and well-known in a general sense. As far as spatial studies go – in a wide range of scientific domains, mainly between Geography and History – to read a map is to interpret the territory there represented. In fact, the various definitions of maps, throughout time, state this direct relation between what is being represented (in different scales, with different symbols and signs) and the territory itself. This can even produce overlapping situations between the limits of Geography and Cartography. Therefore, territorial comprehensions are object of map-making. But how can we consider the different relations established between cartographical productions and what subsidises this territorial knowledge? In this subject, how can textual sources operate in order to create new territorial comprehensions that, ultimately, instigate alternative cartographic elaborations? This presentation will thus cover possible relations that can be established between textual sources, cartographic and territorial studies in general, as of the case study of Auguste de Saint-Hilaire, a french naturalist that traveled the early years of the Empire of Brazil (1816-1822). How can techniques that operate within the Digital Humanities context offer new visualisations on the study of a historical-geographical subject, permitting different approaches and interpretations with the analysis of a textual source, as a journey account? The production of travel maps are, naturally, a first solution for this problem and they hold a great prominence in this domain, but what other visualisations are possible, stimulating renewed territorial interpretations of historical textual sources?

### **BIBLIOGRAFIA**

- [1] AMICO R. (1995), L'archivio del Corpo degli Ingegneri d'Acque e Strade del Compartimento di Pisa, Rassegna degli Archivi di Stato, Roma.
- [2] BARSANTI D. (1985), "Per una 'rilettura' delle forme dell'assetto territoriale del Pisano nei secoli XVI-XIX: la schedatura delle geopiante dell'Ufficio Fiumi e Fossi di Pisa", in *Geografia*, 1, pp.42-51.
- [3] BERTIN, Jacques (1967). *Semiologie graphique: Les diagrammes, les réseaux, les cartes*. Paris: Mouton.
- [4] BIAGIOLI G. (1975), *L'agricoltura e la popolazione in Toscana all'inizio dell'Ottocento. L'indagine sul catasto particellare*, Pacini Editore, Pisa.
- [5] BIAGIOLI G. (1987), *Analisi di alcune fonti toscane: Catasti, Statistiche, Censimenti*, in *Villes et territoire pendant. La période napoléonienne. France et Italie*, Ecole Française, Roma.
- [6] BIAGIOLI G. (2008,) *Gli ultimi Lorena in Toscana: la cultura del territorio*, in *Sovrani nel giardino d'Europa. Pisa e i Lorena*, Pacini Editore, Pisa.
- [7] BONIN, Serge, (2000). *Le développement de la graphique de 1967 à 1997*. Cybergeo: European Journal of Geography.

- [8] CACIAGLI C., CASTIGLIA R. (2001), I disegni degli ingegneri della Camera di Soprintendenza Comunitativa di Pisa, Ministero per i beni e le attività culturali Direzione Generale per gli Archivi, Stamperia e Legatoria Pisana, Pisa.
- [9] CRESPO SOLANAS, Ana (2014). Spatio-temporal Narratives: Historical Gis and the Study of Global Trading Networks (1500-1800). Cambridge Scholars Publishing.
- [10] CRESTI C., ZANGHERI R. (1978), Architetti e Ingegneri nella Toscana dell'Ottocento, Polistampa, Firenze.
- [11] D'Asburgo Lorena P. L. (1969-1974), Relazioni sul governo della Toscana, a cura di A. Salvestrini, Olschki, 3 voll, Firenze.
- [12] DIAZ F. (1997), I Lorena in Toscana. La Reggenza, Utet, Torino.
- [13] Dragicevic S. (2004), "The potential of Web-based GIS", Geographical Systems, Vol. 6, pp. 79-81.
- [14] GENET J.-Ph. (1994), Source, métasource, texte, histoire, in Storia & Multimedia, Grafis Edizioni, Bologna.
- [15] Giuntini A. (1989), La formazione didattica e il ruolo nell'amministrazione granducale dell'ingegnere della Toscana di Leopoldo II, in Ciuffoletti Z. e Rombai L., La Toscana dei Lorena. Riforme, territorio, società, Olschki, Firenze.
- [16] GREGORY, I., ELL, P.S. (2007) [Historical GIS: Technologies, methodologies and scholarship](#). Cambridge: Cambridge University Press.
- [17] GUARDUCCI A. (2009), L'utopia del catasto nella Toscana di Pietro Leopoldo. La questione dell'estimo geometrico-particellare nella seconda metà del Settecento, All'Insegna del Giglio, Firenze.
- [18] INGHIRAMI G., De' Ricci L. (1834), Relazione finale al granduca della deputazione sopra il catasto, Firenze.
- [19] Itzcovich O. (1993), L'uso del calcolatore in storiografia, Franco Angeli, Milano.
- [20] INGOLD, Tim, (2016). Lines: A Brief History. Routledge.
- [21] KRAAK, MJ. (2004), "The role of the map in a Web-GIS environment", Journal of Geographical Systems, Vol VI, pp. 83-94.
- [22] LÜNEN, Alexander von; TRAVIS, Charles (2013). History and GIS: epistemologies, considerations and reflections. Dordrecht: Springer.
- [23] MACCHI Jánica G. (2003), "Ricerca storica e geografia quantitativa", Trame nello spazio. Quaderni di geografia storica e quantitativa, 1, Siena, pp. 11-25.
- [24] MANETTI A. (1848), Alcune parole sulla istituzione del Corpo degli Ingegneri di Acque e Strade, specialmente per ciò che concerne il servizio delle Comunità, s.i.t., Firenze.
- [25] Nuti L. (1986), Pisa. Progetto e città. 1814-1845, Pacini, Pisa.
- [26] MORETTI, Franco, (2003). Atlas do romance europeu: 1800-1900. São Paulo: Boitempo.
- [27] MORETTI, Franco, (2013). Distant reading. London ; New York: Verso.
- [28] MORETTI, Franco, (2007). Graphs, maps, trees: abstract models for literary history. London; New York: Verso.
- [29] PANATTONI R. (2004), Ridolfo Castinelli (1791-1859): architetto e ingegnere negli anni del Risorgimento. Progetti e realizzazioni per committenti privati, Pacini, Pisa.
- [30] PANZIERI M. (2009), Storiografia digitale e metafonti per la storia del territorio tra specificità disciplinari, standard web e dinamiche della rete, in Fonti, metafonti e GIS per l'indagine della struttura storica del territorio, Celid, Torino.
- [31] ROMBAI L., TOCCAFONDI D., VIVOLI C. (1987°), "Cartografia e ricerca storica, un problema aperto. I fondi cartografici dell'Archivio di Stato di Firenze", in Società e Storia, 36.
- [32] RUMSEY D. (2002), Williams M., Historical Maps in GIS, Past Time, Past Place: GIS for History, Esri Press.

- [33] TOCCHINI L. (1961), “Usi civici e beni comunali nelle riforme leopoldine”, *Studi Storici*, Anno 2, No. 2, pp. 223-266.
- [34] TUFTE, Edward R, (2006). *Beautiful evidence*. Cheshire: Graphics Press.
- [35] TUFTE, Edward R (1995). *Envisioning information*. Cheshire.: Graphics Press.
- [36] TUFTE, Edward R (2001). *The visual display of quantitative information*. 2nd ed. Cheshire: Graphics Press.
- [37] TUFTE, Edward R (1997). *Visual explanations: images and quantities, evidence and narrative*. Cheshire: Graphics Press.
- [38] VALENCIA, C, (2016). Precisión y exactitud en los sistemas de información geográfica (sig) en las investigaciones históricas. In: VALENCIA, Carlos & GIL, Tiago. *O retorno dos mapas. Sistemas de informação Geográfica em História*. Porto Alegre: Ladeira Livros.
- [39] VALENCIA, Carlos & GIL, Tiago, (2016). *O retorno dos mapas. Sistemas de informação Geográfica em História*. Porto Alegre: Ladeira Livros.
- [40] VIVOLI C. (2000), “I fondi cartografici e la ricerca storica: il caso toscano”, in *L’Universo*, LXXX, pp. 848-865.
- [41] ZANGHERI R. (1973), *I catasti*, in *Storia d'Italia*, vol. IV, I documenti, tomo I, Einaudi, Torino.
- [42] Zorzi A. (2000), “Documenti, archivi digitali, metafonti”, *Archivi & Computer. Automazione e beni culturali*, X, pp. 274- 291.

# **WORKSHOPS**

# Visualizing the Italian Literary Canon through Distant Reading

Tiziana Mancinelli<sup>1</sup>, Simone Rebora<sup>2</sup>, Rachele Sprugnoli<sup>3</sup>

Università di Venezia<sup>1</sup>, Georg-August-Universität Göttingen<sup>2</sup>, Fondazione Bruno Kessler<sup>3</sup>  
tiziana.mancinelli@gmail.com, simone.rebora@phil.uni-goettingen.de,  
rachele.sprugnoli@alumni.unitn.it

## ABSTRACT

The tutorial will provide an overview of the major methods and tools for the computational analysis of literary texts. The main goal of the tutorial will be that of providing participants with the theoretical and practical knowledge that is necessary for accomplishing a distant reading of texts. In particular, participants will be guided through the steps that connect analysis with visualization, i.e. the automated transformation of “plain text” into annotated, structured, and statistical data, that can be interpreted through visualizations. There will be a specific focus on stylometry, semantic annotation, and network analysis.

All analyses will be run on the Liber Liber corpus of Italian literature, showing how these techniques can be adopted for both didactic and research purposes, e.g. by showing the stylistic similarities between authors and genres in a literary canon, or the direct correlation between a story and the geographical space in which it takes place.

## KEYWORDS

Data visualization, distant reading, stylometry, network analysis, semantic annotation.

## TUTORIAL DESCRIPTION

This half-day tutorial will be divided in three main parts with the aim of providing both theoretical background and hands-on experiences of the analysis and visualization of literary texts. To this end, we will prepare a corpus of Italian literary works in txt format taken from the Liber Liber digital library<sup>1</sup> to be used by the participants.

In the first part, we will briefly introduce the audience to the main themes of the tutorial explaining what stylometry, semantic annotation and distant reading are and why they can play an important role when applied to the literary field. The idea is to discuss both the potential and the limitations of these approaches as applied to the Italian literary canon, by encouraging participants’ active contributions.

In the second part of the tutorial we will focus on stylometry, i.e. the computational study of authorial style. The first section of this part will be dedicated to the use of the Stylo R package<sup>2</sup> [1]. An experiment will be run on a sample corpus of Italian literary texts, with the aim of showing the main features and visualizations that can be used in stylometry. The second section will be dedicated to stylometric visualizations of larger corpora with the Gephi platform<sup>3</sup> [2].

The third part of the tutorial will be devoted to the semantic annotation of a text taken from our corpus using the Recogito<sup>4</sup> web-based platform [3]. Participants will learn how to perform geotagging and georesolution of place names taking advantage of the gazetteers made

---

<sup>1</sup> <https://www.liberliber.it/online/>

<sup>2</sup> <https://github.com/computationalstylistics/style>

<sup>3</sup> <https://gephi.org/>

<sup>4</sup> <https://recogito.pelagios.org/>

available through the tool. We will also make use of the new functionality of Recogito for the annotation of relations between pairs of entities. The final annotated text will then be shared in order to enrich the Linked Open Data web [4].

We have prepared an online questionnaire so that we have a good idea in advance of the competences of the participants<sup>5</sup>, in order to adapt the tutorial to the needs of those taking part. We plan to organize the tutorial so that it can be followed by people with different skills.

## RELEVANCE OF THE TUTORIAL

This tutorial offers training to anyone with an interest in applying Digital Humanities methods to traditional subjects and teaching activities, including academics at all career stages and anyone trying to create an interdisciplinary digital humanities class.

Technology is used daily by students in a passive way, without having any knowledge of how the information they are using is built and shaped. Introducing Digital Humanities' methods into a traditional class could spark new insights into well-known texts, and allow students to explore the function of technology together with features and aspects of literature, thus contextualizing the technological moment in which they live [5].

We will use some masterpieces of Italian literature as a corpus on which to practice basic Digital Humanities methodologies and tools for distant reading, including visualizations, thus showing how texts can be explored in many different ways. The challenge is to face such a wide array of practices in a simple way and has multiple benefits. It introduces participants to open source tools and resources that are freely accessible online: it will shed light on the importance of digital libraries and their analysis through several computational methods. This will also encourage participants to rethink their definition of books and of reading practices.

## TUTORS

Tiziana Mancinelli is currently doing a postdoc at the University of Venice in the ERC project BIFLOW (Bilingualism in Florentine and Tuscan Works, ca. 1260 – ca. 1416), working on Francesco Barberino's medieval manuscript and catalogue. She is also collaborating with the CCeH (Cologne Center for eHumanities), managing the technical coordination, programming, server administration, and data curation for their *Magica Levantiva* project. After her first degree in Italian Studies at the University of Rome, *La Sapienza*, where she was inspired by the pioneers of Digital Humanities, she collaborated on some of the very first Digital Humanities projects. She holds a PhD in Italian Literature and Digital Humanities from the University of Reading (UK), where she completed a dissertation on a scholarly digital edition of 'La camera da letto', by Attilio Bertolucci.

Simone Rebora works as a research fellow between the University of Verona and the Göttingen Centre for Digital Humanities. He holds a BSc in Electronic Engineering (Polytechnic University of Torino) and a PhD in Foreign Literatures and Literary Studies (University of Verona). He is a member of the International Society for the Empirical Study of Literature and Media (IGEL) and of the Distant Reading for European Literary History COST Action. His main research interests are the theory and methods of literary historiography and reader response studies. In the field of digital humanities, he focused on tools and methods such as OCR, stylometry, and sentiment analysis. His essays have appeared in journals such as *Modern Language Notes* and *Digital Scholarship in the Humanities*.

---

<sup>5</sup> <https://goo.gl/forms/2r6PXvRWr3SliGLz1>

Rachele Sprugnoli is a researcher in the Digital Humanities group at the Fondazione Bruno Kessler (FBK) in Trento. She obtained her bachelor and master degrees in Humanities Computing at the University of Pisa and has a PhD in Information Technology from the University of Trento. Since 2005 she has been involved various Human Language Technologies projects working mainly on the semantic annotation of texts, uses of crowdsourcing techniques for Natural Language Processing (NLP) tasks, and the evaluation of NLP tools. In the last five years her research has focused on how computational methods and technologies can be applied to the treatment of cultural content, in particular that of historical texts.

## **TECHNICAL SUPPORT REQUIRED**

- Projector;
- Wifi connection;
- We will ask the participants to bring their own PCs.

## **TUTORIAL LOGISTICS**

Please fill this questionnaire in order to let us know useful information to better tailor the workshop to your needs:

<https://goo.gl/forms/AQR8r5ShTP5Es35E3>

Registered participants will be able to download the material of the tutorial and instructors will send instructions to install software.

The tutorial will be held in Italian or in English depending on the participants who will register to it.

## **REFERENCES**

- [1] Eder, Maciej, Mike Kestemont, and Jan Rybicki. 2016. "Stylometry with R: A Package for Computational Text Analysis." *The R Journal* 8 (1): 1–15.
- [2] Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy. 2009. "Gephi: An Open Source Software for Exploring and Manipulating Networks." *International AAAI Conference on Web and Social Media; Third International AAAI Conference on Weblogs and Social Media*. <https://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>.
- [3] Simon, Rainer, Elton Barker, Leif Isaksen, and Pau de Soto Cañamares. "Linked Data Annotation Without the Pointy Brackets: Introducing Recogito 2." *Journal of Map & Geography Libraries* 13, no. 1 (2017): 111-132.
- [4] Ciotti, F., Lana, M. and Tomasi, F., 2014. TEI, ontologies, linked open data: Geolat and beyond. *Journal of the Text encoding Initiative*, (8).
- [5] McGann, Jerome. *Radiant textuality: literary studies after the World Wide Web*. Springer, 2016.



# **LONG PAPERS**

# Per una filologia (digitale) degli scartafacci

Elena Pierazzo<sup>1</sup>, Alessia Marini<sup>2</sup>

<sup>1</sup>Université Grenoble-Alpes, Francia, elena.pierazzo@univ-grenoble-alpes.fr

<sup>2</sup>Università di Siena, Italia, alessia\_marini@outlook.it

## ABSTRACT

La realizzazione di un prototipo di edizione digitale di alcuni autografi manzoniani fornisce materiale di riflessione sullo statuto filologico e ontologico degli scartafacci d'autore. Le autrici si interrogano inoltre sull'apporto del digitale nello studio dei manoscritti d'autore.

## KEYWORDS

filologia genetica, testualità, filologia d'autore

I manoscritti d'autore (o scartafacci, per usare la terminologia resa celebre da Contini) stanno gradualmente diventando oggetti privilegiati della riflessione filologica.

Soprattutto in territorio francese, essi non sono visti solo come forieri di un'opera in divenire che, teleologicamente, si svilupperà in un testo artisticamente compiuto, ma in sé, come oggetti di studio capaci di rivelare il processo autoriale oltre a rendere manifeste le motivazioni artistiche dell'autore. L'aumento della visibilità e dell'attenzione della critica verso questi oggetti è una diretta conseguenza dell'avvento del digitale, che da un lato grazie alle riproduzioni facsimilari ha permesso a un gran numero di documenti del passato di diventare vere e proprie star di internet, dall'altro fornisce strumenti di analisi e di comprensione adatti sia al lavoro di ricerca che alla divulgazione scientifica. Il prototipo relativo ai manoscritti proustiani realizzato nel 2012 (André e Pierazzo 2013; Pierazzo 2014) è un caso tipico in questo senso: nato grazie alla disponibilità in linea dei facsimili delle carte proustiane realizzati dalla Bibliothèque Nationale de France e dalla volontà di esplorare carte ai fini di ricerca, il prototipo si è rivelato un ottimo strumento anche per la divulgazione, per far capire “come lavorava Proust”.

Gli strumenti digitali a disposizione non sono però semplicissimi da realizzare: richiedono infatti competenze filologiche, di codifica, di trasformazione, di rappresentazione e di creatività; per esempio, il prototipo proustiano, nella sua semplicità, richiede competenze abbastanza avanzate di XML-TEI, SVG, XSLT, XPath, HTML, CSS e JavaScript. In realtà, però, la difficoltà maggiore è stata rappresentata dalla creazione del modello di codifica, che a sua volta riflette un'analisi del manoscritto da un punto di vista filologico e critico. Uno dei problemi principali nella modellizzazione degli scartafacci è dato dal fatto che nonostante sul manoscritto d'autore ci sia *del* testo, non è sempre detto che ci sia *un* testo, anzi: secondo Daniel Ferrer uno scartafaccio è solo una ricetta per creare un testo (Ferrer 1998). Il problema si complica quando le ricette si moltiplicano e lo scartafaccio contiene gli ingredienti di molti testi. È il caso del manoscritto del Fermo e Lucia / Sposi Promessi. Il manoscritto autografo che ospita la cosiddetta Prima Minuta (Manzoni 2006), è servito infatti da base al Manzoni per l'elaborazione della Seconda Minuta, conosciuta anche con il titolo di Sposi Promessi (Manzoni 2012). Ciascuna della carte del primo tomo del romanzo conservate alla Braidense, presenta sulla colonna di destra la stesura del Fermo e Lucia (FL) e su quella di sinistra gli Sposi Promessi (SP), con il *caveat* che a volte alcuni frammenti della colonna di destra (FL) sono, per così dire, “promossi” in SP, e hanno quindi una doppia valenza e una doppia testualità mentre alcuni frammenti della colonna di sinistra risalgono alla stesura di FL.

Questa complessa situazione documentaria ma soprattutto testuale è alla base di un esperimento/prototipo digitale che è stato realizzato in risposta a una sfida lanciata da Giulia Raboni e Paola Italia, nell'ambito del progetto "Manzoni Online: carte, libri, edizioni, strumenti (PRIN 2015FN4ZSN)". Tale progetto ha come obiettivo la messa in linea di versioni affidabili e editorialmente curate delle opere manzoniane, oltre alla produzione dei facsimili degli autografi, delle prime edizioni, delle lettere e dei postillati d'autore. Si affianca a questo obiettivo principale quello di produrre edizioni pilota/laboratorio di alcune carte particolarmente complesse e stratificate.

La sfida menzionata prima è quella di verificare la capacità dello schema TEI a gestire la doppia testualità presente nel documento, oltre che la sua natura composita: non sono infatti rare le pagine che includono cartigli e foglietti incollati attestanti ulteriori strati di scrittura oltre ai due menzionati. La modellizzazione del testo si presentava particolarmente complessa con molteplici obiettivi:

- 1) Esplorare la complessità della pagina autografa;
- 2) Ricostruire le versioni del Fermo e Lucia e degli Sposi Promessi attestata dal manoscritto in tutte le loro stratificazioni;
- 3) Dare conto del processo autoriale che si sviluppa nel corso di sette anni;
- 4) Evidenziare quali parti del manoscritto appartengono a entrambi i testi in modo da dare conto del metodo di lavoro e del processo creativo del Manzoni.

La codifica adottata si basa sull'elemento TEI <sourceDoc> il quale permette di includere la trascrizione del testo all'interno di una gerarchia documentaria, fatta di superfici e zone invece che di sezioni e paragrafi, come nello schema TEI più tradizionale. Invece di limitarci a distinguere i due testi, abbiamo concepito il modello su un asse temporale, cercando di riprodurre con la codifica il processo di scrittura: prima abbiamo delimitato le zone con il testo del solo FL, poi abbiamo ripercorso il processo mentale manzoniano da un lato aggiungendo le zone contenenti il testo di SP, dall'altro attribuendo un'etichetta alle zone già esistenti del FL, dichiarando se esse erano da considerarsi *cancellande* oppure erano da mantenersi. L'atto della codifica è diventato così un momento privilegiato dell'analisi filologica.

Il prototipo realizzato consente di visualizzare i due processi di scrittura (FL e SP), di estrarre i testi dell'uno e dell'altro "in pulito", e di evidenziare le sovrapposizioni fra i due. L'obiettivo è quello quindi di mettere in luce la stratificazione della testualità manzoniana in un modo che, secondo noi, risulta più semplice e intuitivo di quanto non facciano le edizioni critiche con i relativi apparati che sommano la difficoltà della lettura del testo dovuta alla grande quantità di segni diacritici, al fatto che sono pubblicate in volumi separati, spezzando quindi quella relazione feconda fra i due che invece è presente nel manoscritto e che caratterizza il lavoro manzoniano.

Di là dei risultati ecdotici sul testo manzoniano, l'esperimento pone la base per ulteriori riflessioni metodologiche sulla natura e sul futuro della filologia degli scartafacci, come momento privilegiato di riflessione sul testo, sul documento e sul farsi dell'opera letteraria.

## REFERENCES

- [1] André, Julie, e Elena Pierazzo. 2013. «Le codage en TEI des brouillons de Proust: vers l'édition numérique». *Genesis* 36 (13): 155–61.
- [2] Ferrer, Daniel. 1998. «The Open Space of the Draft Page: James Joyce and Modern Manuscripts». In *The Iconic Page in Manuscript, Print and Digital Culture*, a cura di Georges Bornstein e Theresa Tinkle, 249–267. University of Michigan Press. <http://www.item.ens.fr/index.php?id=23616>.
- [3] Manzoni, Alessandro. 2006. *I Promessi Sposi (1821-1827). Edizione critica diretta da Dante Isella. 1.*

*Fermo e Lucia. Prima minuta (1821-1823)*. A cura di Barbara Colli, Paola Italia, e Giulia Raboni. Vol. 1. Milano: Casa del Manzoni.

[4] ———. 2012. *I Promessi Sposi (1821-1827). Edizione critica diretta da Dante Isella. 2. Gli Sposi Promessi. Seconda minuta (1823-1827)*. A cura di Barbara Colli e Giulia Raboni. Vol. 2. Milano: Casa del Manzoni.

[5] Pierazzo, Elena. 2014. «Unpacking the draft page: a new framework for digital editions of draft manuscripts». *Variants* 11: 29–46.

# Can we do better than computers? Teaching Computational Thinking to Digital Humanists

Marilena Daquino<sup>1</sup>, Silvio Peroni<sup>2</sup>, Francesca Tomasi<sup>3</sup>

<sup>1</sup> Department of Classic Philology and Italian Studies – University of Bologna, Italy – marilena.daquino2@unibo.it

<sup>2</sup> Department of Classic Philology and Italian Studies – University of Bologna, Italy – silvio.peroni@unibo.it

<sup>3</sup> Department of Classic Philology and Italian Studies – University of Bologna, Italy – francesca.tomasi@unibo.it

## ABSTRACT

In this paper we present the course “Computational Thinking and Programming” available at the Digital Humanities and Digital Knowledge (DHDK) International MA of the University of Bologna. The course is designed to offer Digital Humanists both theoretical lectures and practical hands-on with a strong focus on problem solving in the Digital Humanities and Library and Information Science. *Learning-by-example* and *collaborative learning* are applied to foster teamwork and accomplish a given task. Finally, the *design-science methodology* is applied to evaluate learning outcomes.

## KEYWORDS

Digital Humanities; Computational thinking; teaching practices

## INTRODUCTION

Creating computational models, constructs, and methods allows us to solve problems that we would not be able to address otherwise. When designing innovative artefacts we test their suitability with regard to an intended purpose, we try to solve a problem, and we learn about the real world (Hevner et al. 2004). The achievement of such objectives requires a number of computational skills.

Computational thinking (CT) has been defined as an universally applicable attitude that is not exclusively adopted by computer scientists. “Computational thinking involves solving problems, designing systems, and understanding human behaviour, by drawing on the concepts fundamental to computer science.” (Wing 2006, 33). It includes the ability to think iteratively and recursively, abstract and decompose a complex problem, design large systems, use invariants to describe the systems, apply heuristic reasoning, and so on. To this extent, computational thinking goes far beyond the artefacts that can be produced by using such a skill. It encompasses the computational concepts used to approach a task.

Digital Humanities methodologies often embrace the *design-science method*, i.e. the production of digital artefacts to cope with some issue, as a means (1) to extend human capabilities in the networked information age and (2) to expand “the reach and relevance of the humanities far beyond small groups of specialists locked in hermetically sealed conversation” (Burdick et al. 2012, 82). However, despite technology is highly linked to DH methodologies, these do not overlap. DH scholars formulate their research questions so as that “their solutions can be represented as computational steps and algorithms” (Aho 2012). In other terms, they generally embrace CT to formalise their questions and to achieve their goals, regardless the technological output.

How to develop such computational skills as a necessary background is subject of several works. In the 1960s, Alan Perlis argued that students of all disciplines should learn

programming and the “theory of computation” (Guzdial, 2008). Grover and Pea (2013) surveyed the state of art on CT in K-12 education, highlighting gaps in curricula for nurturing computational competencies, such as the underinvestigated idea of computing as a medium for teaching other subjects. In this paper we present the course “Computational thinking and Programming”<sup>1</sup> available at the Digital Humanities and Digital Knowledge (DHDK) MA of the University of Bologna. The course is designed to foster CT applied to DH common and sophisticated research tasks.

In the following sections we describe existing CT courses and their scope (Related works), we introduce the objectives, the structure and the methods underpinning the design of the course (Computational Thinking and Programming for Digital Humanists), the methodologies applied to teach and learn CT (Methodologies for learning CT). Finally, we conclude with some considerations on the evaluation of the course and future improvements.

## RELATED WORKS

Both Academy and Industry offer courses, both online and offline, addressing CT. The Open University offers an online course called “Introduction to Computational Thinking”,<sup>2</sup> which briefly addresses some applications of CT in biology, physics, economics and sport. Humanities related topics are not taken into account, and no other applications are shown than an overview of Machine Learning techniques. The Vrije University Amsterdam in collaboration with Huygens Instituut and the KNAW Humanities Cluster in Amsterdam offers a more sophisticated CT course<sup>3</sup> devoted to students with different backgrounds, including humanities, social science and computer science. The course is part of the minor Digital Humanities and Social Analytics. Courses for educators are available too. For instance, Google online course “Computational Thinking for Educators”<sup>4</sup> is targeted at high school teachers and provides an introductory-level overview of topics related to CT.

## COMPUTATIONAL THINKING AND PROGRAMMING FOR DIGITAL HUMANISTS

The Digital Humanities and Digital Knowledge (DHDK) International MA<sup>5</sup> started in Bologna in 2017, directed by Professor Francesca Tomasi. Around 20 students were selected and enrolled at the beginning of the first academic year, and about 15 students attended classes of Computational Thinking course and finalised the exam. Among these, the majority had a humanities background and had no prior knowledge of programming or CT, while a few others had a Computer Science background.

The objectives of the course can be summarised as follows:

- To know the high-level principles, as well as the historical and theoretical backgrounds, for solving problems efficiently by using computational tools and information-processing agents.
- Being able to understand and use the main data structures for organising information.
- To develop algorithms for addressing computational-related tasks, and implement such algorithms in a specific programming language (i.e. Python).

---

<sup>1</sup> <http://www.lettere.beniculturali.unibo.it/it/corsi/insegnamenti/insegnamento/2017/424624>

<sup>2</sup> <http://www.open.edu/openlearn/science-maths-technology/computing-and-ict/introduction-computational-thinking/content-section-0?active-tab=description-tab>

<sup>3</sup> <https://www.vu.nl/en/study-guide/2018-2019/minor/d-e/digital-humanities-and-social-analytics/index.aspx?view=module&origin=50968570&id=50049587>

<sup>4</sup> <https://computationalthinkingcourse.withgoogle.com/>

<sup>5</sup> <https://corsi.unibo.it/2cycle/DigitalHumanitiesKnowledge>

Students were asked to attend 30-hour face-to-face classes, held by Silvio Peroni (Assistant Professor in Computer Science), and 16-hour hands-on classes, held by Marilena Daquino (Research Assistant in Digital Humanities). Fifteen face-to-face classes (two hour each) provided the theoretical background on a number of CT related topics, such as algorithmic techniques and methods, computability, information structures, and recursion. Every lesson included insights on the historical background, biographical notes on notable scientists relevant to the topic at hand, and the definition of CT concepts, supported by the usage of Python-like code so that students acquainted with the programming language since the beginning. Six practical sessions were planned so as to leverage the knowledge acquired during the theoretical classes and implement algorithms in Python. Python was chosen because it is a widely-used, readable, high-level programming language, that can be potentially reused in different situations (e.g. prototyping, data mining, text analysis, web development) that a DH scholar may encounter during her research path.

The final exam included a written test (divided in three sections: theory, understanding and practising) and the presentation of group project. The project required groups to develop an efficient and effective application to solve a task given by the teacher. In particular, the last year task involved the creation of a Python class implementing a text-based search engine for querying bibliographic data. The project was designed to cope with real world problems that can be encountered in DH and Library and Information science researches.

All the contents of the course are available on two bespoke github repositories, that include lectures and practical hands-on.<sup>6</sup> Several open source tools have been used to create and share contents of the course. Theoretical lectures are available in HTML, and were originally edited by using the RASH Framework<sup>7</sup> (Peroni et al., 2017), using the RAJE editor<sup>8</sup> (Spinaci et al., 2017). Laboratorial lectures and exercises are issued as Jupiter notebooks,<sup>9</sup> which allow to create and share documents that contain live code, equations, visualizations and narrative text (written in markdown syntax<sup>10</sup>). Jupiter was chosen so as to boost the reproducibility of developed solutions, encourage good practises such as documenting code, and introduce user-friendly markup languages. An online instance of Jupiter was provided to students for creating their own diaries during classes.

## METHODOLOGIES FOR LEARNING CT

CT implies learning how to efficiently cope with a given problem. Despite problem-solving is often deemed a concern for scientists only, Digital Humanities offer plenty of relevant domain-dependent cases. Therefore, practical exercises solved by students were devoted to tackle DH problems in a *learning-by-example* fashion. In particular, four scenarios were addressed, namely:

- Manipulate structured bibliographic metadata so as to query, sort, cleanse, count, and reorganise data (e.g. get all the authors, order references by author, remove duplicates from a list, transform references according to a given citation style).
- Analyse a corpus of unstructured texts by leveraging basic Natural Language Processing tools and extract meaningful information on a pattern (e.g. detect recurring expressions in a text that characterise a music listening experience).
- Reconcile entities included in authority files to records belonging to well-known online datasets (e.g. DBpedia<sup>11</sup>). Being this the most complex task, the “use–modify–

---

<sup>6</sup> <https://github.com/essepuntato/comp-think> and [https://github.com/marilenadaquino/computational\\_thinking](https://github.com/marilenadaquino/computational_thinking)

<sup>7</sup> <https://rash-framework.github.io>

<sup>8</sup> <https://github.com/rash-framework/raje-app>

<sup>9</sup> <http://jupyter.org/>

<sup>10</sup> <https://daringfireball.net/projects/markdown/>

<sup>11</sup> <https://dbpedia.org>

create” progression is adopted to help the learner go from user to modifier to creator of computational artefact (Lee et al., 2011). Specifically, the student runs and modifies an existing script in order to achieve a slightly different goal.

- Organise bibliographic references based on a Library classification system (i.e. Dewey), extract the hierarchy of concepts, reproduce the arrangement of books on shelfmarks of a library (i.e. the British Library).

To achieve their goals, students were encouraged to adopt *collaborative learning*, *collaborative sense-making*, and *collaborative problem-solving* approaches. Students were grouped to review given information and instructions, to look for potential existing solutions (either online or provided by other students), and to work together in realising the technical solution.

Lastly, the design-science methodology was (partially) applied as a learning method by asking students to develop an outstanding solution for accomplishing the final exam. The evaluation of the produced algorithms was performed by the teacher, who tested iteratively who among the students was able to deliver the fastest solution.

## CONCLUSIONS AND FUTURE WORKS

In this paper we presented the course “Computational Thinking and Programming” available at the Digital Humanities and Digital Knowledge (DHDK) International MA of the University of Bologna. It provides a set of tools, methods and practices for pragmatically teaching CT to Digital Humanists, by leveraging real world scenarios and well-known learning and research methodologies. The course was held only one time, hence we are not able at this stage to provide a complete evaluation. Future improvements of the course include the adoption of test-driven development as agile development approach to propose to students for the creation and test of their algorithms and programs, and a reorganisation of the scheduling of theoretical and practical classes, which will be alternated so as to formalise a fixed knowledge retention period and measure its efficiency by means of a user-satisfaction evaluation. All the outcomes of the evaluation questionnaires filled up by the students will be analysed by using grounded theory techniques (Strauss and Corbin, 1998).

## REFERENCES

- [1] Aho, Alfred V. "Computation and computational thinking." *The Computer Journal* 55.7 (2012): 832-835.
- [2] Burdick, Anne, et al. *Digital\_Humanities*. Mit Press, 2012.
- [3] Hevner, Alan, and Samir Chatterjee. "Design science research in information systems." *Design research in information systems*. Springer, Boston, MA, 2010. 9-22.
- [4] Lee, Irene, et al. "Computational thinking for youth in practice." *Acm Inroads* 2.1 (2011): 32-37.
- [5] Wing, Jeannette M. "Computational thinking." *Communications of the ACM* 49.3 (2006): 33-35.
- [6] Peroni, Silvio, et al. "Research Articles in Simplified HTML: a Web-first format for HTML-based scholarly articles." *PeerJ Computer Science* 3 (2017): e132.
- [7] Spinaci, Gianmarco, et al. "The RASH JavaScript Editor (RAJE): A Wordprocessor for Writing Web-first Scholarly Articles." *Proceedings of DocEng 2017* (2017): 85-94.
- [8] Strauss, Anselm, and Corbin, Juliet. *Basics of Qualitative Research Techniques and Procedures for Developing Grounded Theory (2nd edition)*. Sage Publications, 1998.



# Developing digital research workflows among undergraduates

Mike Cosgrave

University College Cork, IE, m.cosgrave@ucc.ie

## INTRODUCTION

Much of the literature approaches interdisciplinarity from the perspective of the problems faced by experienced researchers moving beyond disciplinary silos. At that level, emphasis is placed on reconciling the wide range of different possible epistemological positions. Here, I consider the issue from the perspective of teaching the much narrower range of research methods to undergraduates across a range of fields, and most recently in an interdisciplinary digital humanities program.

Many studies on inter- and transdisciplinarity treat the discipline as the basic unit of study, and seek to map new bonds between disciplines. Disciplinary methods are seen as a key part of the package which makes up a discipline. However, while disciplines have different ontological and epistemological perspectives, the actual range of methods through which those are applied to the sources is quite limited. In Digital Humanities, a focus on digital tool building demands that one look inside the disciplines to discern what are the fundamental methods of the disciplines, and how can those methods be expressed as digital processes.

## EXPLORING DISCIPLINES

Creating tools to give digital effect to the methods of various disciplines exposes similarities between methods which individual disciplines have treated as unique to their field.

Years of teaching undergraduate research methods courses within disciplines like History, International Relations or Politics, even where those methods have become increasingly digital is much simpler than teaching research methods in an explicitly interdisciplinary field like Digital Humanities. Most disciplines have working, if often debated, definitions which can be presented on a single powerpoint slide but introducing eighteen year olds to an interdisciplinary programme brings in several extra steps when speaking to an audience who have never had to consider the nature of a discipline, much less what lies in the “swampy lowlands” between disciplines (Schön, Donald A. "Knowing-in-action: The new scholarship requires a new epistemology." *Change: The Magazine of Higher Learning* 27, no. 6 (1995): 27-34.)

My own research and teaching experience has been quite ‘undisciplined’. While I was always interested in military history, my first graduate work was the history of a local brewery from 1792 to the 1970s, a project which required exploring not only business history and accounts, but also the economic history context, the social history of drinking in Ireland, the history of the families who ran the brewery, including their activities as local philanthropists in the city, and their participation in local and national politics. The conclusions were unsurprising, but the range of sources and methods were broad. I was fortunate to continue my research and teaching in a school of history that was always broadly ecumenical in its approach, covering not only conventional history, but also contemporary politics and international relations. As a result I taught not only undergraduate survey courses in history and digital history to our own final year undergraduates, but also at various times economic history to business students, game theory to international relations postgraduates and research methods to politics undergraduates. I did not therefore follow the ‘normal’ pattern of ever deeper immersion in a

narrow field, I always preferred wide perspectives and this has given me a different perspective on methods.

Being thrust into delivering a new undergraduate programme in Digital Humanities which explicitly aims to be inter and transdisciplinary in bridging the gap between the Humanities and IT required explaining interdisciplinarity to undergraduates who had not previously been required to understand what disciplinarity meant. The need to equip digital humanities undergraduates with the ability to use digital tools to apply humanities research methods to the cycle of research, analysis and writing demands clarity about what we mean by interdisciplinarity.

## **TOOLS AND METHODS**

For the digital humanist, tools and methods may provide a point of attack on the problem of understanding interdisciplinarity. Humanities disciplines often take pride in eschewing explicit methods, but in practice they do have signature methods, and those methods are often similar. The ‘digital’ demands that we model work habits and methods explicitly in order to build the computational tools to apply our methods to the materials related to the problem being explored. In so doing, it exposes common elements in humanities methods. When the working methods of various humanities fields, locked in silos and often treated more as art than science, are exposed to view, they are not so unique or precious as disciplinarians imagine.

### *Disciplinary Methods*

Scholars who have grown up with the security blanket of a major ‘old’ discipline tend to uncritically assume their field is not only important, distinctive and unique in both subject and methods, but also eternal. This is actually incorrect: until the Enlightenment, the model of a scholar was the ‘Renaissance man’, versed in all fields of human knowledge. Disciplines only emerge in the 1700’s when knowledge becomes too diverse for the ‘everyman’; and only become formal when they acquire university chairs and budgets.

Thus, while history traces its intellectual origins back to Herodotus and Thucydides, as a modern discipline it was a nineteenth century political project, created to support the nation state, exploring the history of the state through official “primary” sources. While history has changed greatly, its core method remains “reading the sources” to understand why certain choices were made.

Wineburg has explored how the active source reading of researchers and students exposes purposeful human choices by linking the sources to their context. Historians tend not to physically annotate sources, but using a read aloud protocol, Wineburg has shown how students, teachers and academics use their contextual knowledge to annotate the sources, connecting them to a wider understanding of the frames of reference of historical figures to explain their actions (Wineburg, Samuel S. "On the reading of historical texts: Notes on the breach between school and academy." *American Educational Research Journal* 28, no. 3 (1991): 495-519. See also Wineburg, Sam

“Historical Thinking and Other Unnatural Acts: Charting the Future of Teaching of teaching the past” Temple University Press, 2001) . It this sort of exploration of of world views brings historians into the realm of phenomenological epistemology.

Annotating text is more explicitly visible in literary studies. Whether you believe the author is dead or alive, the signature pedagogical artifact in literary studies across every language for

learners from secondary school to the grave is the annotated text, the scholarly edition. Layers of notes explain not only archaic language, metaphors and images, but also how these are associated in the intertextual space to create or embody meanings. ( Allen, Graham. *Intertextuality*. Routledge, 2011)

Annotating for technique and meaning is common to all of subfields of art history, although here as in history the researchers make stand off annotations rather than annotating the ‘text’. Moving quickly to a wider view, we might ask: when researchers in this broad area venture out to write notes on works in galleries, how exactly do these notebooks differ from Darwin’s notebooks full of observations of natural life? Or an ethnographers field notes from a Brazilian tribe or a teenagers in a shopping mall? All of these are annotating something in order to discover patterns, in pursuit of either inductive, deductive or abductive answers. In this they are in some ways encoding what they observe, and this glides gently into the social sciences. Interviews, focus groups and survey results are all in various ways reduced to patterns by the application of codes. Whereas in the humanities the coding is loose and unstructured, inductive coding, the coding of social science research is more organized and explicit. Textbooks like Pierce ( see Chapters 17, 18 & 19 of Pierce, Roger *Research Methods in Politics: A Practical Guide*, Sage, 2008) ) demonstrate this process, using the interviews of Kenneth Harris to move through the process from interviews to coding to visualizations to reduce the reflections of political elites to codes and patterns for comparison. (Harris, Kenneth, *Conversations*, Hodder & Stoughton, London, 1967, especially pp 255-286)

A key part of reading the texts, whether in humanities or social sciences, is locating the source and its creator in their network of influences. Artists, poets, musicians and politicians are all the products of schools which influenced their development, as well as colleagues with whom they debated. Traditional methods struggle to adequately map the network of influences on individual creators. Addressing this, social network analysis has become a common part of Digital Humanities projects. The set of public sample datasets for Gephi, the dominant tool in this area, covers topics from the co-appearance of characters in *Les Miserables*, through visualising the relationship between developers of the PERL programming language, the network of relationships between characters in the Marvel comics universe to the network topology of the western US and word adjacencies in *David Copperfield* - in other words, the method, expressed in this particular tool, is applied across a great many disciplines. ( Datasets, <https://github.com/gephi/gephi/wiki/Datasets> Accessed 30 Jan 2018) What is important here is that one method dominates questions of this type across the whole range of disciplines. Visualization and network analysis have become key parts of the methods toolkit of most Humanities and Social Sciences inquiries.

The changing content of methods textbooks is also illustrative here. In the past, the most sophisticated advice in humanities methods texts was to take notes on index cards for sorting by theme or topic as patterns ‘emerged’ from the sources: indeed, one could argue that history was the archetype of ‘grounded theory’ before that method ever had a name. Social science methods texts originally dealt exclusively with quantitative methods, mainly survey based research. While there are now many texts dealing with specific types of methods - text analysis, surveys, qualitative data, network analysis, action research, ethnographic research and more - synoptic undergraduate research methods texts now feature both qualitative and quantitative methods and speak to the importance of using mixed methods to triangulate research. The use of a number of different methods to verify research not only improves the reliability of findings, but also brings together methods which had previously been seen as the

preserve of particular disciplines. This makes visible how the diverse range of humanities and social sciences rely on a narrow range of methods.

### **Scholarly Primitives and Methodological Commons**

When humanities research was done by hand, methods within disciplines were opaque. Anyone with the time to observe humanities scholars at work would see nothing beyond slow, close reading, furrowed brows and the occasional, often illegible, note. Transferring this to the digital realm, beginning with Fr. Busa's Index Thomisticus, and early work on stylometrics and quantitative history, required making these processes explicit in code to build the tools to apply them.

While the the number of disciplines, fields and specialisms in the Humanities & Social Sciences is broad and continues to grow, the *types* of data studied and the *range* of methods applied to them are actually limited. It is important to distinguish low level methods from higher level methodologies or epistemologies, which are much more diverse. However, just as as certain methods like surveys or electron microscopy can be used in a wide range of epistemological paradigms in the sciences, the handful of methods common to the humanities are used across a range of fields. Driven by the desire to build useful tools in the humanities, John Unsworth (2000) argued that there seemed to him to be a limited number of 'scholarly primitives' which he argued were "activities are basic to scholarship across eras and across media" (2000). He identified these as

- Discovering
- Sampling/Selection
- Annotating
- Comparing
- Linking
- Illustrating
- Representing

Unsworth finished a discussion of these with a demonstration of the possibilities the idea of scholarly primitives by taking a pre-meeting text for a meeting on the Human Genome Project, and editing it to a fictional Humanities Genres Project by changing only "'Humanities Genres Project' for 'Human Genome Project' and 'humanist' for 'biologist' and 'library' for 'laboratory,'" . The resulting text shows how easily many mechanical elements of research organisation transfer easily between STEM and Humanities. In a related paper in 2002, Willard McCarty and Harold Short argued for the existence of a methodological commons in the humanities. At the heart of their argument are five major data type found in the humanities:

- “(1) narrative or "running" text;
- (2) tabular alphanumeric ("chunky") data,
- (3) numbers,
- (4) visual images and
- (5) music, including other kinds of sound.”

(McCarty & Short, 2002, <https://eadh.org/publications/mapping-field> Accessed 22nd Jan 2018. Versions of the diagram are to be found in Siemens, R. (2016). Communities of practice, the methodological commons, and digital self-determination

in the Humanities. *Digital Studies/Le champ numérique*. DOI: <http://doi.org/10.16995/dscn.31> as well as Anderson, Dunn & Blanke, 2010).

These are mapped to disciplinary groups and broad areas of learning. McCarty and Short list only some methods : “text-analysis, database design, numerical analysis, imaging, music information retrieval and so forth. The list is not fixed nor stable, of course, as from time to time new methods are devised.” (McCarty & Short 2002)

McCarthy & Short’ categories of data, matched with Unsworth’s ‘Scholarly primitives’ produces a broadly useful framework of the types of source humanities scholars work with and the methods applied to those sources.

	Narrative Text	Numbers	Tabular alpha- numerics	Music (including other sound)	Images
Discovering					
Sampling/Selection					
Annotating					
Comparing					
Linking					
Illustrating					
Representing					

## Tools v2

If we use this framework to look at the range of work represented in recent Digital Humanities conferences and syllabi, it is difficult to find work which cannot be accommodated in it. Thus even recent work which pushes the limits of technology can be mapped to this framework. Looking at the digital tools mentioned in conference papers or syllabi, we find not only the same methods but also the same tools mentioned again and again. Tools like Gephi, R, MALLET, Voyant, CATMA, Nvivo or similar are used across a wide range of disciplines to apply Unsworth’s ‘Scholarly Primitives’ to McCarty’s common data types.

Many parts of this matrix are well populated with tools created by the digital humanities community, and some of those tools are even user friendly. Digitization of sources, with improved data curation strategies and metadata standards, make the discovery and selection of digital data for research led teaching much easier. While we now have several standards for annotation of sources, none can compare with the tools which the social sciences use to encode all forms of data: several social science tools allow for easy and consistent tagging of conceptual patterns across all types of source, but, as they are proprietary and non-standards compliant, they fail when it comes to exporting results in standard formats to represent the results. In theory, digital humanities should be able to make use of the “notebook” style tools now becoming common in the sciences which include text, data and code in one package, but in practice we do not.

Where there is a weakness, it is in the lack of a workflow. While there are digital tools which express all the parts of the common humanities set of methods, as yet there is no tool which covers the full spectrum of activities in an integrated package. As a result, a great deal of time is still spent wrangling the outputs of one stage into the inputs for the next. This, in

interdisciplinary work, distracts from higher level discussions of epistemology which is where disciplinary differences emerge most distinctly.

### **What are Disciplines, really?**

If a great deal of the work in many disciplines in Humanities & Social Sciences is in practice, done by an identifiable set of methods which could be expressed in a set of digital tools, then one is compelled to ask what is special or unique about disciplines at all?

Many of the established disciplines only emerged since the 1800s, and have changed over the past 150 years.

#### *Disciplines as Evolving Communities*

Disciplines are not only recent in origin, and still evolving, but in the contemporary university, while disciplines are strictly defined in administrative terms, in terms of scholarship, they vary considerably.

#### *Limits to Disciplines*

It seems that there are certain limitations on disciplines as effective units. Firstly, disciplines are successful when the community of inquiry is of a manageable size. Small disciplines wither, while larger ones tend to splinter into specialized subfields. The significant growth in the number of university academics since the 1960's has made it difficult to maintain contact, both personal and intellectual, with all the developments in the 'old' disciplines - where old often means only a 100 years old or less!

Secondly, well defined disciplines worked well when knowledge had a longer lifespan. As both Kuhn and Latour demonstrated, the pace at which paradigms shift is now faster. It is certainly shorter than the career span of an academic, and this has serious implications for career paths and for disciplinary management. The old process was that heads of discipline would recruit and promote researchers who followed in their footsteps but a head of discipline might now have difficulty finding intellectual disciples or keeping pace with emerging issues. Thirdly, disciplines worked well when problems could be separated neatly and studied in isolation. With the emergence in succession of systems theory, complexity theory, chaos theory and holistic world views like General Systems Theory, or the Gaia hypothesis, it is clear that problems can no longer be neatly isolated in disciplinary "clean rooms" for study. The acronym VUCA, short for volatile, unpredictable, complex and ambiguous, is becoming common shorthand for the sort of conditions and research problems which we face now and into the future.

### **When does interdisciplinary first appear?**

Interdisciplinarity has been advanced as a solution to these limits. It is now also the case that the preponderance of research funding is for interdisciplinary projects - indeed the proportionate reduction in available funding for focussed, disciplinary research is potentially a serious a problem.

Indeed, recently the calls for completely reorganising the university on interdisciplinary grounds have been advanced. Writing in the *Guardian*, Zahir Irani announced that at Bradford:

"We're removing departmental divisions and restructuring ourselves around research. Under this approach, research centres— based around interdisciplinary expertise and collaborations – administer taught courses, using research to inform course creation and delivery. The structure is intended to encourage cooperation between staff and

students, strengthen the ties between teaching and research activities, and turn collaborative, interdisciplinary working into the norm”

(Irani, Zahir, The university of the future will be interdisciplinary” Guardian ONline, Accessed 30th Jan 2018 <https://www.theguardian.com/higher-education-network/2018/jan/24/the-university-of-the-future-will-be-interdisciplinary>

This proposal is not surprising in the light to the growth of the ‘interdisciplinarity industry’ but it may provoke a hostile reaction and start a debate with more heat than light.

A more nuanced view on the development and relationships of disciplines allows one to see that, whatever about administrative structures and silos, intellectually disciplines have been in a process of development over the past century. It has been argued here that disciplines have significant commonality of methods on which to build new relationships. In terms of producing ‘work ready’ graduates, certain tools and methods have wide applicability. There are differing views on how far one can introduce undergraduates to epistemology and ontology, there are a limited set of methods through which students may be led to develop a grasp of different ways of framing questions using the same tools.

Making explicit the key disciplinary methods is also tackled from the perspectives of student learning and pedagogical design. Thus the Decoding the Disciplines project (<http://decodingthedisiplines.org/> ) interrogated academics about their teaching and assessment models while the Visible Knowledge Project (<http://gallery.carnegiefoundation.org/collections/exhibits/vkp/index.htm>) at Georgetown explored student learning in detail in order to make visible what and how students actually learn. As with the creation of digital scholarly methods and workflows, projects like Decoding the Disciplines and VKP which make learning visible are as important for the development of research based learning at undergraduate level (see Healey, Mick. "Linking research and teaching exploring disciplinary spaces and the role of inquiry-based learning." *Reshaping the university: new relationships between research, scholarship and teaching* (2005): 67-78. For an entry point into this topic) as work on developing digital scholarly infrastructures and workflows being done by projects like CLARIN and DARIAH

Methods are like cooking: there are only four great sauces in French cooking, but in the hands of a master, there are infinite variations. A good cocktail needs only three or four ingredients (a spirit, a sweet, a sour and a twist). There are a limited palette of methods, but an endless range of epistemological terroir. However, understanding the tools and methods which disciplines] have in common can provide a practical basis for interdisciplinary collaboration. Gathering all this together would permit a more organic, evolutionary way of seeing the development of scholarly fields, rather than a confrontation between discipline and interdiscipline.

# **L'ecosistema digitale per la cultura del piemonte. Strumenti e modelli per la condivisione della conoscenza dei beni culturali archivistici, librari e museali**

**Dimitri Brunetti**

Regione Piemonte, [dimitri.brunetti@regione.piemonte.it](mailto:dimitri.brunetti@regione.piemonte.it)

## **ABSTRACT**

L'intervento si propone di illustrare il percorso realizzato dalla Regione Piemonte per la condivisione libera e gratuita del patrimonio di conoscenze riferito ai beni culturali archivistici, librari e museali prodotto nel corso degli anni grazie al lavoro di numerosi soggetti pubblici, istituti culturali e operatori della cultura con interventi di descrizione, catalogazione e digitalizzazione. In particolare, il contributo vuole soffermarsi sui modelli concettuali, sui set di regole e sulle realizzazioni web per la produzione di descrizioni e oggetti digitali nonché per la loro gestione in rete all'interno di un articolato ecosistema digitale per la cultura.

## **Favorire la pubblica fruizione e la valorizzazione dei beni culturali**

La Regione Piemonte, in ambito culturale, ha il compito di promuovere la conoscenza, la valorizzazione e la fruizione dei patrimoni dei musei, delle biblioteche e degli archivi dei soggetti pubblici e privati, in quanto testimonianza ed espressione della cultura e della storia dei territori, delle comunità e delle persone. Deve quindi promuovere e sostenere interventi di salvaguardia, conservazione, descrizione, digitalizzazione, pubblicazione e valorizzazione; deve favorire la creazione e lo sviluppo di reti, sistemi e altre opportune forme di cooperazione sul territorio.

Per realizzare i propri obiettivi di coordinamento e dare corpo alla propria *mission* quale promotore di occasioni di crescita e di condivisione, la Regione sviluppa strumenti innovativi per la descrizione dei beni culturali, la raccolta di oggetti digitali, la gestione dei dati e delle informazioni e l'esposizione sul web. Inoltre, intervenendo sugli strumenti di lavoro si favorisce la costituzione e il consolidamento di comunità professionali, l'accumulo di patrimoni informativi di qualità rispettosi degli standard, la gestione dei materiali anche da parte di enti e soggetti culturali che non riescono ad adottare sistemi proprietari e, inoltre, si garantisce l'accesso alle informazioni e alla rappresentazione dei beni sia sul territorio regionale, sia su quello nazionale e oltre.

Entro il contesto descritto, il progetto Mèmore si propone di dare corpo alla visione in tema di sistemi informativi, con la finalità di descrivere le diverse tipologie di beni culturali nel pieno rispetto degli standard e delle regole, di promuovere la conoscenza del patrimonio culturale piemontese, la ricerca e la collaborazione tra enti e cittadini, di informare il pubblico, i ricercatori, gli enti e le istituzioni, di facilitare interventi di salvaguardia e valorizzazione del patrimonio.

La sfida è quella di sviluppare un nuovo concetto di sistema che permetta di definire i progetti e le aggregazioni, di descrivere in modo integrato i beni culturali multitematici, di ordinarli, organizzarli e validarli, di stabilire molteplici relazioni fra le risorse e di integrare gli oggetti digitali con le risorse. Una piattaforma a disposizione dei titolari dei beni archivistici e museali e degli operatori, ma orientata alla pubblicazione e alla condivisione della conoscenza.



## Collective Access e Mèmora

Negli anni scorsi la Regione Piemonte e il CSI-Piemonte hanno dato avvio allo sviluppo dell'Ecosistema digitale dei beni culturali (in sostituzione dei precedenti applicativi denominati Guarini e Guarini Patrimonio Culturale), un sistema integrato di applicativi per la descrizione dei beni culturali, la raccolta di oggetti digitali correlati, la gestione dei dati e delle informazioni e l'esposizione sul web sia per un pubblico professionale, sia per un'utenza generalista. Un Ecosistema che si propone come aggregatore unitario dei progetti già realizzati, di quelli in corso e di quelli che si svilupperanno nei prossimi anni.

Mèmora è l'applicativo open source piemontese per la descrizione, la gestione e la pubblicazione delle informazioni sul patrimonio culturale archivistico e museale. L'attuale configurazione è il risultato dello sviluppo di moduli e plugin che hanno dato una nuova forma al software Collective Access. Il sistema offre la possibilità di schedare il patrimonio archivistico e quello di ambito museale sia per tipologia di bene, sia in modo integrato, così da rappresentare al meglio anche le specificità e la ricchezza di fondi multitematici, in cui coesistono beni culturali di diversa natura.

Collective Access è un software multilingua e multiplatforma con licenza open source GPL che permette di descrivere, organizzare e gestire documenti e raccolte di musei, biblioteche e archivi. La componente di catalogazione di base applica gli standard internazionali di descrizione catalografica, archivistica e biblioteconomica (ISAD, ISAAR, DublinCore, DACS, MARC, PBCore), consente l'implementazione di nuovi standard e integra set di dati personalizzati. Il sistema prevede la possibilità di creare profili diversi per l'accesso e la visualizzazione delle risorse. È possibile completare i dati descrittivi con materiali multimediali (immagini, audio, video, documenti digitali), geolocalizzare il dato e stabilire relazioni tra i diversi elementi del sistema. Garantisce l'interoperabilità e la condivisione delle risorse con altri sistemi informativi poiché i dati possono essere esportati in XML-EAD e XML-EAC.

Mèmora, che è finalizzato alla fruizione delle informazioni e degli oggetti digitali resi disponibili dagli operatori, associa ai contenuti la licenza open source CC-BY-NC-SA, identificata come quella più idonea anche in relazione alle istanze di alcuni soggetti aderenti più prudenti nel condividere liberamente. Questa licenza consente la condivisione e la modifica dei materiali originali a condizione di riconoscere una menzione di paternità adeguata, di segnalare l'avvenuta modifica, di non utilizzare i materiali per scopi commerciali e di distribuirli con la stessa licenza. In realtà la Regione suggerisce ai partner titolari dei diritti di aderire alle politiche regionali sugli open data e adottare la licenza CC BY 4.0, in piena sintonia con l'orientamento nazionale.

Mèmora è l'espressione di un progetto regionale realizzato dal CSI-Piemonte e rientra nelle politiche autonome dell'ente, però un percorso di questo genere deve essere condiviso e sostenuto dal confronto con i soggetti che a vario titolo hanno competenza sulla materia. Così si è cercato di consolidare relazioni e di avviarne di nuove, nonché di sostenere il tradizionale dialogo con gli enti e gli istituti. Il progetto è alla base di convenzioni e accordi di collaborazione stipulati con l'Istituto di informatica e telematica del CNR di Pisa e con l'Istituto per il lessico intellettuale europeo e storia delle idee (Iliesi) del CNR di Roma; con l'Istituto centrale per gli archivi – ICAR; con l'Istituto centrale per il catalogo unico – ICCU. Sono in corso contatti e scambi di documenti con l'Ufficio nazionale per i beni culturali ecclesiastici della Conferenza episcopale italiana e con l'Istituto centrale per il catalogo e la documentazione – ICCD.

Nel luglio 2018 sono state pubblicate le *Linee guida per i progetti di digitalizzazione*, definite di concerto con l'ICCU, che costituiscono un complesso di regole e raccomandazioni per la realizzazione di progetti di digitalizzazione. Il documento illustra i formati e le risoluzioni degli oggetti digitali per immagini, audio, video e 3D, definisce i metadati da utilizzare e la loro struttura, fornisce indicazioni sulla denominazione dei file, sull'organizzazione delle cartelle e sulla consegna dei materiali.

## L'Ecosistema: Mèmora, Librinlinea e Giornali Locali

Mèmora si configura come il primo elemento un percorso più complesso, ovvero l'ecosistema digitale per la cultura. Mèmora, che è autosufficiente per le fasi di lavoro, espone i suoi dati verso Yucca - Smart Data Platform, la piattaforma cloud della Regione Piemonte abilitante l'ecosistema per la valorizzazione dei dati, in grado di gestire i contenuti in modo svincolato

dalle griglie di partenza e di rimodularli nei modi più opportuni per dialogare con altre piattaforme. Entro il 2019 ci si propone di condividere i dati con i portali nazionali SAN, Sigec, Cultura Italia e Internet Culturale, ma anche con Europeana grazie alla collaborazione con l'ICCU, ed eventualmente con altri portali e altre piattaforme. Grazie all'interoperabilità OAI-PMH e all'esposizione dei dati di Yucca sarà possibile per i soggetti che partecipano all'ecosistema regionale gestire i propri dati ed esporli sul sito dell'ente o istituto in modo da offrire agli utenti l'intero catalogo dei propri beni. Yucca, coerentemente con le indicazioni Agid e del *Piano triennale per l'informatica nella pubblica amministrazione*, oltre a supportare la creazione di nuovi servizi legati al tema della valorizzazione dei dati, costituisce il repository degli opendata di titolarità della Regione e degli enti che aderiscono alle politiche regionali in materia, integrandosi già ad oggi con le piattaforme nazionali.

Il front-end di Mèmora, già realizzato ma non ancora pubblicato, in prospettiva, sarà in grado di rendere fruibili contestualmente ai dati relativi agli archivi e al patrimonio culturale anche i materiali residenti in due altri sistemi regionali: [www.giornalidelpiemonte.it](http://www.giornalidelpiemonte.it) e [www.librinlinea.it](http://www.librinlinea.it).

I giornali locali sono un patrimonio prezioso per la storia del territorio e costituiscono un bene di grande interesse dal punto di vista sociale, economico, culturale e storiografico. Le informazioni quotidiane sono un elemento di straordinaria importanza per conoscere cosa succede nel nostro presente, ma sapere anche cosa è accaduto in passato è altrettanto significativo per ricostruire i frammenti della nostra identità. Inoltre, i giornali locali sono l'unica fonte in grado di raccontare giorno per giorno le vicende dei nostri territori, delle comunità e delle persone.

Purtroppo trovare notizie su di un giornale di carta non è affatto semplice: bisogna avere a disposizione le testate e i fascicoli e bisogna cercare con grande attenzione avendo fortuna. Così, per valorizzare l'informazione giornalistica locale e per facilitare la ricerca è stato realizzato un nuovo Portale, unico in Italia e in continua crescita.

Il portale Giornali del Piemonte è stato realizzato dal Consiglio e dalla Giunta Regionale, con la collaborazione della Federazione italiana piccoli editori giornali – FIPEG, degli editori che hanno condiviso questo progetto, degli istituti culturali, dei Sistemi bibliotecari e delle biblioteche pubbliche.

Il Portale digitale dell'informazione giornalistica piemontese trae origine dalle numerose campagne di digitalizzazione realizzate dai Sistemi bibliotecari e dalle biblioteche pubbliche già a partire dalla fine degli anni Novanta. Il portale si avvale dell'esperienza del progetto di digitalizzazione del quotidiano "La Stampa", recupera le testate storiche già disponibili sulla Teca digitale piemontese e di altre piattaforme. L'obiettivo è quello di rendere disponibili online la copia dell'intera raccolta di tutti i giornali locali pubblicati in Piemonte.

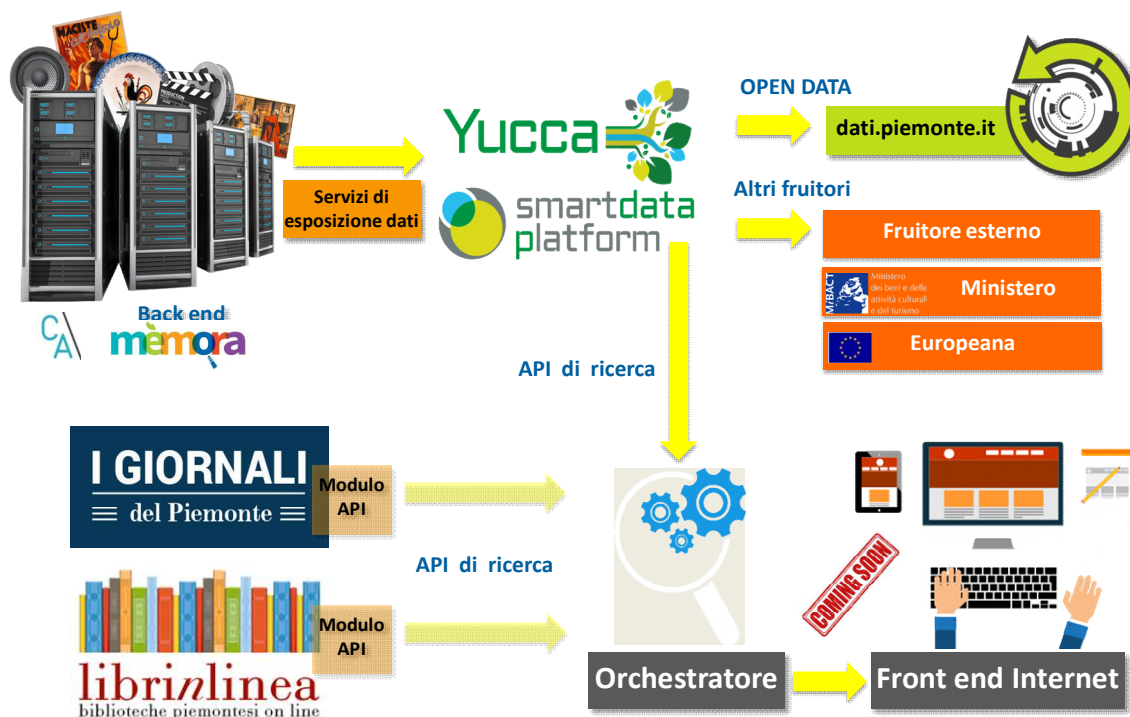
Nel marzo 2016 è stato reso pubblico il portale [www.giornalidelpiemonte.it](http://www.giornalidelpiemonte.it) che nell'ottobre 2018 è stato completamente rinnovato rendendo disponibili altri periodici, oltre ai giornali locali, riferiti a temi e ambiti professionali specifici raccolti in apposite pagine tematiche. Oggi ai cittadini, ai ricercatori, agli insegnanti, agli studenti e ai giornalisti sono disponibili gratuitamente oltre tremilioni e duecentomila pagine di più di duecento testate di giornali e periodici a partire dal 1846.

Librinlinea è l'OPAC che consente l'accesso e l'interrogazione alle risorse documentarie delle biblioteche, dei sistemi e delle reti che hanno aderito al progetto con l'obiettivo di diventare un meta-catalogo on line attraverso il quale è possibile consultare in maniera integrata le risorse disponibili nelle biblioteche di tutto il Piemonte, indipendentemente dal Polo del Servizio Bibliotecario Nazionale cui esse partecipano.

Con gli archivi e i musei provenienti da Mèmora, le biblioteche da Librinlinea e i giornali si ricompongono l'unità dei luoghi e degli istituti della cultura e si restituisce una visione armonica

del patrimonio cui si accede attraverso il front-end quale elemento essenziale e terminale dell'ecosistema digitale per la cultura e il turismo in Piemonte.

Il Portale di esposizione e di interrogazione intende valorizzare la ricchezza e la qualità del patrimonio culturale piemontese, la varietà dei luoghi di produzione e di conservazione e soprattutto gli istituti che contribuiscono a creare e condividere nuova conoscenza. Il front-end si proporrà come principale punto di accesso al patrimonio culturale piemontese, offrendo percorsi e linee temporali, focus su ambiti e temi specifici e agevoli modalità di ricerca, georeferenziazione e modalità alternative di esplorazione, garantendo un'esperienza d'uso in mobilità, con interfacce adattabili a molteplici strumenti e con la vocazione al multilinguismo.



**Figura 1. Architettura dell'Ecosistema digitale per la cultura e il turismo in Piemonte.**

## BIBLIOGRAFIA E SITOGRAFIA

- [1] Brunetti, D. 2016, *La lente archivistica: per rendere convergenti percorsi catalografici paralleli. Appunti sulla multidisciplinarietà della descrizione*, in «Archivi» Rivista dell'ANAI, n. 1/2016, pp. 101-114.
- [2] Brunetti, D. 2014, *Testate giornalistiche d'informazione locale: interventi di digitalizzazione e valorizzazione in Piemonte*, in «DigItalia» 1/2014, pp. 89-101.
- [3] Brunetti, D. 2018, *L'archivio multimedializzato: definizione, descrizione e identità*, in «La Gazette des archives», periodico dell'Associazione degli archivisti francesi, n. 1/2018 (249), pp. 39-49
- [4] Brunetti, D. 2018, *Mèmore. La nuova piattaforma digitale per i beni culturali piemontesi*, in «DigItalia» 1/2018 (in corso di stampa)
- [5] [www.memora.piemonte.it](http://www.memora.piemonte.it)
- [6] [www.librinlinea.it](http://www.librinlinea.it)
- [7] [www.giornalidelpiemonte.it](http://www.giornalidelpiemonte.it)
- [8] [www.collectiveaccess.org](http://www.collectiveaccess.org)



# SPARSAR recites Shakespeare's Sonnets – and Coping with Early Modern English variants

Rodolfo Delmonte

Università Ca Foscari, Italy, delmont@unive.it

## ABSTRACT

In this paper I will present *SPARSAR* - the system for the analysis and reading of English poetry – and the problems I had to cope with in dealing with Early Modern English texts, Shakespeare's Sonnets written in Elizabethan English. I will also explain what consequences ensue from the presence of diachronic variants on the recital by TTS(Text-To-Speech).

Shakespeare wrote the sonnets before his death in 1616 in a period in which Early Modern English was still in use but starting to undergo substantial changes. We don't know precisely how words were pronounced but we know for sure what an Elizabethan sonnet should sound like due to rhyming conventions which were very strict at the time, poetry being mainly rehearsed and only occasionally available on printed paper. In this way, specific words may be deemed to have a different sound from the current one. Since a lexicon of a language should encompass the basic sound of a word, together with its grammatical, syntactic and semantic properties, we will be concerned with the way in which a system for poetry recital through TTS should account for these variations.

## KEYWORDS

NLP, Sentiment and Affective Analysis, Prosodic Structure, Semantic and Syntactic Processing, Metrical Structure, Rhetorical Devices, Rhyming Devices, Rhyming Scheme, Poetry Visualization, TTS Recital.

## INTRODUCTION

The system for English poetry analysis called *SPARSAR* - System for Poetry Automatic Rhythm and Style Analyzer and Reader<sup>1</sup> - has been recently upgraded in order to take care of Elizabethan poetry of the XVIth and XVIIth century. In particular, we have been dealing with the collection of Shakespeare's sonnets which constitute a fairly representative corpus from a statistical point of view. In order to be able to recite a Shakespearean sonnet, the system has undergone an extended number of modifications in all its components.

In a recent work (6), I showed how the union between sound and meaning in poetry is very strong: the sonnets can be verified to communicate sad emotions whenever there's a majority of unvoiced sounds, high or back vowels, obstruent consonants; the contrary applies for those sonnets where the words are meant to communicate positive sensations of happiness, where it is possible to find a majority of voiced sounds, low or middle vowels, sonorant and continuant consonants. *SPARSAR* system automatically computes metrical and rhyming structure. Elizabethan and Shakespearean sonnets are organised on the basis of the iambic pentameter and have alternate rhymes distributed into three quatrains and a final couplet which is in perfect rhyme. All sonnets, excluding no. 99 - which is made up of 15 lines organised in chained rhyme -, and no. 126 which only has 12 lines. All lines are pentameter feet sometimes with feminine ending, i.e. with one additional syllable (11 instead of 10), again excluding no. 145 which is made of lines tetrameter feet. In many cases, however, respecting iambic

---

<sup>1</sup> Presented for the first time at an international conference in 2013(4) as a general analyzer and poetry reader it has now been specialized for the new task.

pentameter is guaranteed by the presence of a great number of contractions, and the opposite prohibition of elision of end of word unstressed syllable. In the following sections I will briefly introduce the system for semantic and phonetic analysis, then I will concentrate on peculiarities of Elizabethan poetry and in particular on how to cope with it. The system will then be demonstrated at the conference by having it analyze and read aloud one sonnet.

### ***SPARSAR*, a System for Poetry Analysis and Reading**

The system can work on any type of poem and produces a set of parameters that are then used to compare poems with one another, of the same author or of different authors. The output can be visualized as a set of colored boxes of different length and width and allows a direct comparison between poems and poets (5). In addition, parameters produced can be used to evaluate best similar candidate poems by different authors by means of Pearson's correlation coefficient. The system uses a modified version of *VENSES*, a semantically oriented NLP pipeline (3). It is accompanied by a module that works at sentence level and produces a whole set of analysis both at quantitative, syntactic and semantic level. The second module is a rule-based system that converts each poem into phonetic characters, it divides words into stressed/unstressed syllables and computes rhyming schemes at line and stanza level. To this end it uses grapheme to phoneme translations made available by different sources, amounting to some 500K entries, and include CMU dictionary, MRC Psycholinguistic Database, Celex Database, plus our own database made of some 20,000 entries. Out of vocabulary words are computed by means of a prosodic parser we implemented in a previous project (1).

The system has no limitation on type of poetic and rhetoric devices, however it is dependent on language, in this case English. Rules implemented can demote or promote word-stress on a certain syllable depending on selected language, line-level syllable length and contextual information. This includes knowledge about a word being part of a dependency structure either as dependent or as head. A peculiar feature of the system is the use of prosodic measures of syllable durations in msec, taken from a database created in a previous project (1). We produce a theoretic prosodic measure for each line and stanza using mean durational values associated to stressed/unstressed syllables. We call this index "prosodic-phonetic density index", because it contains count of phones plus count of theoretic durations: the index is intended to characterize the real speakable and audible consistency of each line of the poem. A statistics is issued at different levels to evaluate distributional properties in terms of standard deviations, skewness and kurtosis. The final output of the system is a parameterized version of the poem which is then read aloud by a TTS system: parameters are generated taking into account all previous analysis including sentiment or affective analysis and discourse structure, with the aim to produce an expressive reading(6).

### **CAPTURING RHYME AND RHYTHM**

As Tsur(9) comments in his introduction to his book, iambic pentameter has to be treated as an abstract pattern and no strict boundary can be established. The majority of famous English poets of the past, while using iambic pentameter have introduced violations, which in some cases – as for Milton's *Paradise Lost* – constitute the majority of verse patterns. Instead, the prosodic nature of the English language needs to be addressed, at first. English is a stress-timed language as opposed to Spanish or Italian which are syllable-timed languages. As a consequence, what really matters in the evaluation of iambic pentameters is the existence of a certain number of beats – 5 in normal cases, but also 4 in deviant ones. Unstressed syllables can number higher, as for instance in the case of exceptional feminine rhyme or double rhyme, which consists of a foot made of a stressed and an unstressed syllable (very common

in Italian), ending the line - this is also used by Greene et al.(8) to loosen the strict iambic model.

In our system, in order to allow for variations in the metrical structure of any line, we operate on the basis of syntactic dependency and have a stress demotion rule to decide whether to demote stress on the basis of contextual information. The rule states that word stress can be demoted in dependents in adjacency with their head, in case they are monosyllabic words. In addition, we also have a promotion rule that promotes function words which require word stress. This applies typically to ambiguously tagged words, like "there", which can be used as expletive pronoun in preverbal position, and be unstressed; but it can also be used as locative adverb, in that case in postverbal position, and be stressed. For all these ambiguous cases, but also for homographs not homophones, tagging and syntactic information is paramount.

Our rule system tries to avoid stress clashes and prohibits sequences of three stressed/three unstressed syllables, unless the line syntactic-semantic structure allow it to be interpreted otherwise. Generally speaking, prepositions and auxiliary verbs may be promoted; articles and pronouns never.

## CONTRACTIONS VS. RHYME SCHEME

Contractions are present in great number in the sonnets. Computing them requires reconstructing their original complete corresponding word form in order to be able to match it to the lexicon or simply derive the lemma through morphological processing. This is essentially due to the fact that they are not predictable and must be analysed individually. Each type of contraction has a different manner to reconstruct the basis wordform. In order to understand and reconstruct it correctly, each contraction must go through recovering of the lemma. We have found 821 contractions in the collection, where 255 are cases of genitive 's, and 167 are cases of past tense/participle 'd. The remaining cases are organised as follows:

- SUFFIXES attached at word end, for example ['s, 'd, 'n, 'st, 't, (putt'st)]<sup>[SEP]</sup>
- PREFIXES elided at word beginning, for example ['fore, 'gainst, 'tis, 'twixt, 'greeting]<sup>[SEP]</sup>
- INFIXES made by consonant elision inside the word [o'er, ne'er, bett'ring, whate'er, sland'ring, whoe'er]<sup>[SEP]</sup> o'ercharg'd, 'rous]

When we completed the analysis of the sonnets (7), we discovered more than 100 violations to the typical sonnet rhyme scheme. This theme has been discussed and reported in papers and also on a website - <http://originalpronunciation.com/> - by the linguist David Crystal(2). His position is however very disputable: he is also convinced that the original pronunciation may be induced by the structure of the rhymes, but then he describes the phenomenon claiming it is merely phonological.

From a careful study of the first 100 words in Crystal's list – which includes also words coming from the plays -, one can easily understand that the phenomenon is not describable simply using phonological rules: the same vowel in stressed position is assigned to an indefinite number of transformation disregarding its context. No reference is made to neighbouring syllables nor to stress patterns. Variants are then lexically determined. To support this approach, we referred ourselves to the book published in 1889 by the famous historical phonologist Wilhelm Viëtor(10).

Some words involved in the transformation are listed below using Arpabet as phonetic alphabet in the excerpt taken from the lexicon. As can be easily noticed, variants are related also to stress position, but also to consonant sounds.

Lexicon 1.

shks(despised,d\_ih2\_s\_p\_ay1\_s\_t,ay1,ay1)

shks(dignity,d\_ih2\_g\_n\_ah\_t\_iy1,iy1,ay1).

shks(gravity,g\_r\_ae2\_v\_ah\_t\_iy1,iy1,ay1).

shks(history,hh\_ih2\_s\_t\_er\_iy1,iy1,ay1).  
shks(injuries,ih2\_n\_jh\_er\_ay1\_z,iy1,iy1).  
shks(jealousy,jh\_ch2\_l\_ah\_s\_iy1,iy1,ay1).  
shks(jollity,jh\_aa2\_l\_t\_iy1,iy1,ay1).  
shks(majesty,m\_ae2\_jh\_ah\_s\_t\_iy1,iy1,ay1).  
shks(memory,m\_ch2\_m\_er\_iy1,iy1,ay1).  
shks(nothing,n\_ah1\_t\_ih\_ng,ah1,ow1).

It is now clear that variants need to interact with information coming from the rhyming algorithm that alone can judge whether the given word, usually at line end - but the word can also be elsewhere -, has to undergo the transformation or not. The lexicon in our case has not been built manually but automatically, by taking into account all rhyming violations and transcribing the pair of word at line end on a file. The algorithm searches couple of words in alternate lines inside the same stanza and whenever the rhyme is not respected it writes the pair in output. Take for instance the pair LOVE/PROVE, in that order in alternate lines within the same stanza: in this case it is the first word that has to be pronounced like the second. The order is decided by the lexicon: LOVE is included in the lexicon with the rule for its transformation, PROVE is not. In some other cases it is the second word that is modified by the first one, as in CRY/JOLLITY, again the criterion for one vs the other choice is determined by the lexicon.

## REFERENCES

- [1] Bacalu C., Delmonte R. 1999. Prosodic Modeling for Syllable Structures from the VESD - Venice English Syllable Database, in Atti 9° Convegno GFS-AIA, Venezia.
- [2] Crystal David, 2009. Sounding Out Shakespeare: Sonnet Rhymes in Original Pronunciation, <http://originalpronunciation.com/>.
- [3] Delmonte R., et al. 2005. VENSES – a Linguistically- Based System for Semantic Evaluation, in J. Quiñero-Candela et al.(eds.), Machine Learning Challenges. LNCS, Springer, Berlin, 344-371.
- [4] Delmonte R. 2013. Computing Poetry Style, in C. Battaglini, C. Bosco, E. Cambria, R. Damiano, V. Patti, P. Rosso (eds.), Proceeding ESSEM - Emotion and Sentiment in Social and Expressive Media: approaches and perspectives from AI (ESSEM 2013), CEUR Workshop Proceedings, Torino, 148- 155, <http://ceur-ws.org/Vol-1096/>.
- [5] Delmonte Rodolfo, 2015. Visualizing Poetry with SPARSAR - Poetic Maps from Poetic Content, Proceedings of NAACL-HLT Fourth Workshop on Computational Linguistics for Literature, Denver, Colorado, June 4, 2015. c 2015 Association for Computational Linguistics, pages 68–78.
- [6] Delmonte Rodolfo, 2016. Expressivity in TTS from Semantics and Pragmatics, in Vayra, M., Avesani, C. & Tamburini F. (Eds.) (2015). Il farsi e disfarsi del linguaggio. Acquisizione, mutamento e destrutturazione della struttura sonora del linguaggio/Language acquisition and language loss. Acquisition, change and disorders of the language sound structure, Milano: AISV. pp. 407-427.
- [7] Delmonte Rodolfo, 2016. Exploring Shakespeare's Sonnets with SPARSAR, in LINGUISTICS AND LITERATURE STUDIES, vol. 4, pp. 61-95; also published online at [http://www.hrpub.org/journals/article\\_info.php?aid=3215](http://www.hrpub.org/journals/article_info.php?aid=3215).
- [8] Greene E., T. Bodrumlu, K. Knight. 2010. Automatic Analysis of Rhythmic Poetry with Applications to Generation and Translation, in Proceedings of the 2010 Conference on EMNLP, 524–533.
- [9] Tsur Reuven. 2012. Poetic Rhythm: Structure and Performance: An Empirical Study in Cognitive Poetics, Sussex Academic Press, 472.
- [10] Vietor Wilhelm, 1909. “A Shakespeare Phonology”, Marburg/London.



# “Digital Autoethnography” and “Connective Intelligence”: Teaching and (Re)Thinking (About) Today’s Techno-Society

Stefano Calzati

Politecnico di Milano, Italy, stefanocalzati@hotmail.com

## ABSTRACT

This article explores the field of Digital Humanities in connection with education. Notably, it does so by adopting a critical/qualitative orientation towards the teaching of new media literacies by providing concrete examples of two practice-based courses.

By now digital natives have reached the stage of undergraduate or postgraduate education. This means that these students have grown up in an increasingly digitalized education system. On the other hand, today’s teachers are still, by and large, members of earlier generations. As such, they represent a cornerstone within our education system in that they have one foot in a pre-digital epoch and the other grounded on today’s technologized society, of which they can detect positive and negative aspects.

Hence, the article discusses the conception, design and results of two practice-based teaching experiences which were aimed at exploring the tensions embedded in our daily use of digital devices and in today’s techno-society as a whole. The first one is a “digital autoethnography” developed, together with Prof. Roberto Simanowski, at the City University of Hong Kong as part of a larger course focused on the processes of self-representation on social networks. The second one refers to the course “Anthropology of Communication”, which I co-delivered with prof. Derrick de Kerckhove at Politecnico of Milan and it adopts a connective intelligence approach that led student to reflect on today and tomorrow’s techno-society.

While the first teaching experience was chiefly a self-reflexive critical approach in connection with the impact of digital technologies on the individual, the second one aimed at mapping the main criticalities within our techno-society and come up with possible solutions. Both courses considered students as active learners/users, insofar as they at the forefront of today digital revolution, but also the subjects most in need of critical tools to face it.

## KEYWORDS

digital autoethnography, connective intelligence, SNSs diet, techno-society, digital natives, new media literacies

## INTRODUCTION

This paper addresses the emerging field of Digital Humanities in connection with education. Notably, it does so by maintaining a critical standpoint towards the impact that digital technologies are having on today’s students and society as a whole. A qualitative practice-based orientation towards the teaching of new media literacies, of which we are increasingly in need at all levels, is advocated. Concrete examples concerning two courses developed by the author, in conjunction with colleagues in Italy and abroad (e.g. Hong Kong), will be provided, in order to discuss the projectuality behind these courses, the research findings coming from them, and the further promotion of teaching experiences in new media literacies (especially by bringing together humanists and computer scientists). Although these two courses are presented together, their underpinning projectuality and objectives are different and the results cannot be comparable.

## STATE OF THE ART

If we take as a starting point the mass diffusion of the Web in the mid-1990 (and mobile phones, although they were not yet smartphones), we see that by now digital natives have reached the stage of undergraduate or postgraduate education. This means that these students have grown up, at least since their first circle in schools, in an increasingly digitalized world, and certainly one in which digital technologies and devices have had a more and more radical impact on their daily life, to the point of deeply affecting also how education is conceived and delivered, as well as how learning is performed. At the same time, digital natives represent the pulling force of digital technology's evolution, meaning that they are both the main target (as consumers) of tech companies and the main producers of digital content and services, providing an epitomizing example of what is meant by the term "producers" (see, among others, Jenkins et al. 2013; van Dijck 2013).

On the other hand, today's teachers and scholars are still, by and large, members of earlier generations, i.e. generations that, to various degrees, have transited from an analogue to a digital society. This means that they were born, raised and educated according to a pedagogical vision that privileged written words over moving images, syntagmatic step-by-step approaches to knowledge over paradigmatic hypertextual ones, and individual reading and memorization over interaction and practice (Hayles 1999). As such, teachers represent a cornerstone within today's education system for motives that go beyond mere pedagogical issues and point, rather, to their bridging role. Indeed, teachers literally keep one foot in an epoch that predates the digital revolution, while the other foot is now solidly grounded on today's technologized society, which they have seen growing since its birth and of which they can pinpoint, for this very reason, both potentialities and shortcomings, notably by retracing its genealogy, its (betrayed and realized) hopes and its actual pervasiveness.

The reason for reinstating what is maybe obvious, although sometimes overlooked – i.e. the encounter in class of two generations that have radically different approaches towards digital technologies – is crucial in order to highlight the fruitfulness of the synergy between today's teachers and students. We need more than ever to rethink teaching and learning as really interactive and mutually beneficial processes for all actors involved, and we can do so by leveraging on the characteristics of these same actors. Students are certainly the subjects who more easily enter in contact and familiarize with (new) digital technologies, precisely because these are conceived for them in the first place. And yet, they largely lack, as one of the two teaching experiences discussed below will show, the tools for making a critical use of technology. Teachers, by contrast, can help students both to put things into perspective – e.g. to investigate the archaeology of new media – and to develop practical skills – e.g. programming – for de-commoditizing technology and make a wise use of it, instead of simply being played by it.

## **“DIGITAL AUTOETHNOGRAPHY” AND “CONNECTIVE INTELLIGENCE”**

With these premises in mind, this paper discusses the conception, design and results of two practice-based teaching experiences which were aimed at exploring, in innovative ways, the tensions embedded in our daily use of digital devices and in today's techno-society as a whole. The first teaching experience relates to a “digital autoethnography” (Ribas and Gajjala 2009) developed at the City University of Hong Kong as part of a larger course focused on the processes of self-representation on social networks (i.e. Facebook and Instagram) and the “datafication of the self” (see Felton 2011) that follows. Specifically, we concentrated on the SNSs profiles of 50 participants who were BA and MA students enrolled in the course “Facebook and Autobiography” that Prof. Roberto Simanowski (2018) and myself purposely designed and taught as part of the research. Apart from traditional lectures, the course was

organized as follows: on the one hand, as analysts we tracked the participants' use of SNSs by entering these platforms and creating a profile that students had to befriend (on a voluntary basis) without however any interventions from our side; on the other hand, we asked our participants to self-reflect upon their SNSs diet (over a period six weeks) through a number of assignments, such as a written diary, two rounds of interviews, and in-class debate, which were all part of the final evaluation for the course. By comparing the insights derived from these two perspectives we managed, on the one hand, to understand better how participants represent themselves on SNSs (indeed, a fragmented representation across different media platforms) and, on the other hand, to sharpen the students' critical awareness concerning their online self-projection. From the materials gathered – diaries, close reading, interviews – we noted that users: 1) often share materials and reply to comments uncritically (i.e. without really checking the content of the posts shared or commented on); 2) forget by and large what they like/share after just a few days. These two phenomena, we suggested, are symptomatic of broader tensions affecting the relation between users and digital technologies, which have notably to do, on the one hand, with the fact that today's information deluge works primarily as a form of distraction (see Adorno 1991; Berardi 2015) and, on the other hand, with the externalization of cognitive faculties – e.g. memory (Stiegler 2010) – which are increasingly outsourced to technology.

The second teaching experience refers to the course “Anthropology of Communication” which I co-delivered during the 2018 fall semester together with prof. Derrick de Kerckhove at Politecnico of Milan. The aim of the course was to map the status of today's techno-society by addressing seven macro-themes: ethics, education, ecology, politics, economy, urban planning, and technology. In order to do so we placed the 54 students enrolled in the course (curriculum in “Design of Communication”) at the centre of the learning experience, being aware that they are, indeed, the pulling force of today's techno-society and the designers of tomorrow's. Specifically, as teachers we provided (apart from traditional lectures) the conceptual framework of the course, which took the form of a cloud of keywords that helped students navigate today's techno-society (the cloud can be reached here: [www.madpress.com/aoc](http://www.madpress.com/aoc); examples or keywords are: “participatory democracy”, “datacracy”, “digital twins”, “trashumanism”, “deep learning”, “smart city”, “social credits”, “transparency”, “net neutrality”). In the first stage, students gathered in groups of three to four and we asked them to research on their keywords in order to provide an in-class up-to-date presentation (on a weekly basis). Subsequently, students clustered in bigger groups composed of six to nine members, according to the affiliation of their keywords to one of the seven macro-themes. The objective was to elaborate a final group project that either outlined the conception of a product or service which addressed one of the criticalities of the afferent macro-theme (such as the unwilling circulation of private data on digital platforms), or defined a chart/manual for good and fair practices in the design/use of technology in connection with the assigned macro-theme. The overall idea was to conceive an emerging sustainable techno-village (similarly to what Google is planning, although under very different corporate-driven conditions, in Toronto). In so doing, students were co-opted as part of a “connective intelligence” approach (de Kerckhove 1997) that aimed at providing key solutions for making tomorrow's techno-society (more) sustainable and inclusive on different (micro and meso) levels and in different areas.

The experience in Hong Kong allowed us (as teachers/researchers) as well as the students (as the main actors of the experience) to have a more “distanced” perception of the daily technologies' diet (and how these use their users). However, given the small cohort of participants and its socio-demographic uniformity (all Asian students between 18 and 22 years old), the findings would require further testing to be confirmed. The experience in Milan, on its part, represented a very proactive approach to the ongoing development of

today's techno-society: students faced and discussed up-to-date issues making the course very pertinent and appealing. And yet, the breadth of the macro-themes likely constituted a limitation to the students' effective conception of their final projects. In both cases, what is important to stress (beyond the research finding) is the projectuality behind the courses, which promoted the interaction teachers-students and students-students and the fostering of new media literacies.

## REFERENCES

- [1] Adorno, W. T. 1991. *The Culture Industry: Selected Essays on Mass Culture*. London: Routledge
- [2] Berardi, F. 2015. *And: Phenomenology of the End*. Cambridge, MA: MIT Press.
- [3] De Kerckhove, D. 1997. *The Skin of Culture*. London: Kogan Page.
- [4] Felton, N. 2011. "Numerical Narratives" *Lecture at UCLA Department Design, Media, Arts in November 15, 2011* (1 hour) (<http://video.dma.ucla.edu/video/nicholas-felton-numerical-narratives/387>)
- [5] Hayles, K. 1999. *How We Became Posthuman*. Chicago: University of Chicago Press.
- [6] Jenkins, H., Ford, H., Greene, J. 2013. *Spreadable Media: Creating Value and Meaning in a Networked Culture*. New York: New York University Press.
- [7] Ribas, N., Gajjala, R. 2007. "Developing Cyberethnographic Research Methods for Understanding Digitally Mediated Identities." *Forum: Qualitative Social Research* 8(3): Art. 35
- [8] Simanowski, R. 2018. *Facebook Society. Losing Ourselves in Sharing Ourselves*. New York: Columbia University Press.
- [9] Stiegler, B. 2010. "Memory." *Critical Terms for Media Studies*, edited by W.J.T. Mitchell and Mark Hansen. Chicago: University of Chicago Press.
- [10] Van Dijck, J. 2013. *The Culture of Connectivity: A Critical History of Social Media*. Oxford: Oxford University Press.

# Oltre la galassia delle Digital Humanities: per la costituzione di una disciplina di Informatica Umanistica

Fabio Ciotti

fabio.ciotti@uniroma2.it

## ABSTRACT

Nel corso dell'ultimo quindicennio il multiforme campo di studi che va sotto il nome di Digital Humanities (DH) è divenuto un fenomeno di primaria importanza sia nel mondo della ricerca e della didattica universitaria (sebbene le differenze tra i vari contesti nazionali siano piuttosto rilevanti), sia nel dibattito culturale. Durante questi anni di espansione, la comunità degli studi di DH globale è divenuta così vasta e diversificata che è difficile individuare al suo interno un fondamento epistemico unitario e condiviso. Nonostante il tema della definizione disciplinare sia allo stesso tempo inflazionato e intrinsecamente ‘difficile’, intendo avanzare un tentativo di sistemazione della questione, con il preciso scopo di rispondere a una delle questioni poste come tema della conferenza AIUCD di questo anno: la collocazione epistemologico-disciplinare delle DH in relazione al sistema dei Settori Scientifico-Disciplinari e del sistema di reclutamento accademico in Italia.

## PAROLE CHIAVE

Definizione Digital Humanities; metodologia DH; fondamenti DH; collocazione istituzionale DH

## INTRODUZIONE

Nel corso dell'ultimo quindicennio il multiforme campo di studi che va sotto il nome di Digital Humanities (DH) è divenuto un fenomeno di primaria importanza sia nel mondo della ricerca e della didattica universitaria (sebbene le differenze tra i vari contesti nazionali siano piuttosto rilevanti), sia nel dibattito culturale. Questa espansione del “campo digitale” nel dominio umanistico ha persino generato una contro-reazione da parte dei cosiddetti umanisti tradizionali. Se fino a qualche anno addietro gli approcci digitali erano ignorati o al massimo considerati con sussiego, ora si succedono prese di posizione critiche e persino allarmistiche, che spaziano dalla critica del loro contenuto scientifico ed epistemologico, a quella della loro valenza politica ritenuta egemonista e dall'impronta “neo-iper-liberista”, (Fish 2012; Marche 2012; Allington, Brouillette, e Golumbia 2016; Tomasin 2017).

Durante questi anni di espansione, la comunità degli studi di DH globale è divenuta così vasta e diversificata che è difficile individuare al suo interno un fondamento epistemico unitario e condiviso. Ne consegue che, se prima dell'espansione era difficile definire le DH come “una” disciplina che condivide oggetto o dominio di studio, assunti epistemologici, metodi e protocolli di ricerca, oggi l'impresa è forse impossibile. Tuttavia il dibattito meta-disciplinare non si esaurisce.

Definire le *Digital Humanities*, è stata una delle preoccupazioni principali che hanno attraversato la comunità, tanto che qualcuno ha parlato dell'esistenza di un vero e proprio genere specifico nella letteratura di settore (Kirschenbaum 2010), genere cui hanno contribuito molti dei suoi esponenti più autorevoli. Basti citare solo alcune delle pubblicazioni uscite negli anni più recenti come la fortunata miscellanea curata da M. Terras, J. Nyhan ed E. Vanhoutte *Defining digital Humanites* (Terras, Nyhan, e Vanhoutte 2013), il libro

*Understanding Digital Humanities* di Berry (Berry 2012), le due raccolte *Debates in Digital Humanities* (Gold 2012; Gold e Klein 2016) e il più recente *Big Digital Humanities* di Patrick Svensson (Svensson 2016), oltre alla produzione di carattere storiografico, che spesso insiste sulla stessa questione (Ciotti 2018).

D'altra parte "andare meta" è un atteggiamento tipico di ogni nuovo campo della ricerca scientifica, specialmente quando esso ambisca ad assumere un ruolo e un riconoscimento nelle istituzioni accademiche, a divenire disciplina in senso istituzionale. Vale la pena far notare che questo non è solo la manifestazione di uno (sterile?) interesse (meta)teorico, come sembra intendere (Ryan 2016), ma anche materialità istituzionale economica ed esistenziale. Si tratta insomma di rispondere a domande e questioni come:

- Dove si fa DH? centri, dipartimenti, laboratori?
- Che cosa si insegna nelle DH? Il problema del syllabo delle DH.
- Chi fa DH? il problema della collocazione e della carriera accademica.
- Cosa è buona DH? il problema della valutazione dei prodotti e dei progetti di ricerca DH.

Per questo, nonostante il tema sia allo stesso tempo inflazionato e intrinsecamente difficile, intendo avanzare un tentativo di sistemazione della questione, con il preciso scopo di rispondere a una delle questioni proposte nella *Call for Papers* della conferenza AIUCD di questo anno: la collocazione epistemologico-disciplinare delle DH in relazione al sistema dei Settori Scientifico-Disciplinari e del sistema di reclutamento accademico in Italia.

## I PARADIGMI DELLE DIGITAL HUMANITIES

Il dibattito sulla definizione disciplinare delle Digital Humanities nell'ultimo ventennio è stato caratterizzato dalla presenza di alcuni paradigmi fondamentali. Il primo e più importante di essi, elaborato sin dalle origini, quando il campo era ancora denominato *Humanities Computing* (o *Informatica Umanistica* in Italia), è il *paradigma metodologico*; come rileva in modo efficace (Kirschenbaum 2010, 55): "At its core, then, digital humanities is more akin to a common methodological outlook than an investment in any one specific set of texts or even technologies". In questo contesto il *principium individuationis* del campo umanistico digitale era costituito da una convergenza metodologica, dalla condivisione di un nucleo di metodi generali come i *methodological commons* - proposti da W. McCarty e H. Short (McCarty e Short 2002) - o le primitive di ricerca - secondo la definizione di J. Unsworth (Unsworth 2000, 2002) - applicabili indifferentemente ai vari ambiti della ricerca in area umanistica.

Nel 2004 con la pubblicazione del famoso *Companion* della Blackwell (Schreibman, Siemens, Unsworth, 2004) il campo di studi umanistico/digitale subisce un vero e proprio cambio di fase e vede, come detto, una espansione fenomenale in termini di addetti, diffusione, status e impatto culturale e sociale. La transizione da *Humanities Computing* a *Digital Humanities* è accompagnata da un parziale superamento del paradigma metodologico, per dare spazio a visioni alternative dei fondamenti del campo di studi, di impianto pragmatista o sociologico. Epitome di tale transizione è il successo della formula "Big Tent Digital Humanities": lanciata come tema principale nella conferenza DH2011 di Stanford, è stata largamente adottata per indicare una visione delle DH come una ampia, diversificata e inclusiva comunità di studiosi che si occupano in vario modo dell'intersezione tra digitale e scienze umane. La formulazione teorica più matura di questa visione è stata data da Ray Siemens, sulla base della nozione di origine antropologica e sociologica di *community of practice* (Siemens 2016): "the notion of the community of practice here offers us a framework

to consider and understand *who* we are via *what* it is we do, *where* we do what we do, and *why* we do it in the way that we do it”.

Sotto molti rispetti correlata a questa visione è la suggestiva e provocatoria tesi pragmatista e strumentalista formulata da S. Ramsey. In un dibattito intervento tenuto nel convegno annuale MLA del 2011, intitolato sintomaticamente *Who's In and Who's Out*, egli individua lo specifico delle DH nella loro natura performativa e applicativa, e di conseguenza nella centralità assunta dalla costruzione di artefatti computazionali (Ramsay 2013, 241): “Personally, I think Digital Humanities is about building things. I’m willing to entertain highly expansive definitions of what it means to build something. I also think the discipline includes and should include people who theorize about building, people who design so that others might build, and those who supervise building. But if you are not making anything, you are not — in my less-than-three-minute opinion — a digital humanist.” Sulla base di questa presa di posizione per un certo periodo in ambito statunitense si è registrata una polemica dialettica tra i pragmatisti (‘smanettoni’, diremmo in italiano), che si riconoscevano nel motto ‘more hack, less yack’, e i Teorici (si noti la T maiuscola), che invece rivendicavano l’importanza di un atteggiamento Teorico-Critico (Nowviskie 2016).

Più recentemente P. Svensson, introducendo la formula *Big Digital Humanities*, ha proposto una visione delle DH basata su una sintesi pluralista tra metodologia, teoria critica e pratica applicativa, e ha individuato nella nozione di “infrastruttura” il suo elemento caratterizzante (Svensson 2016): “In this way, the digital humanities offers an infrastructural and intellectual platform for carrying out work placed between the humanities and the digital. This platform seeks deep connections with humanities disciplines and areas as well as with other fields and initiatives. These multiple epistemic traditions and perspectives contribute to making the digital humanities a dynamic and diverse field. Such curiosity-driven work must be based on respect, intellectual sharpness, and technological innovation. Big digital humanities gets leverage from a combined intellectual, material, and political engagement and can serve as an experimental contact zone for the humanities.”

Svensson, nella sua analisi, con certa eleganza argomentativa, valorizza due ambiti discorsivi che hanno recentemente caratterizzato il dibattito e la ricerca nelle DH, soprattutto in ambiente americano: la convergenza tra la nebulosa teorica della “Teoria critica” e degli studi culturali e gli studi digitali e il riferimento alla funzione politica della ricerca, attraverso una attenzione verso temi come l’inclusione, la diversità, le culture “marginali”, la geopolitica dei saperi.

## **DIGITAL HUMANITIES: UNA GALASSIA INTER- E MULTI-DISCIPLINARE**

Se si guarda oggi al panorama offerto dalle DH a livello globale, insomma, ci si trova di fronte a una molteplicità proteiforme di esperienze e iniziative che spaziano in ogni ambito delle scienze umane e che adottano approcci e metodi anche opposti tra loro. Nemmeno la classica metafora wittgensteiniana delle “somialtanze di famiglia” potrebbe descrivere adeguatamente questa molteplicità, che purttuttavia presenta un proprio profilo identitario.

Usando una metafora astronomica, credo che potremmo descrivere il campo delle Digital Humanities come una galassia. In una galassia, infatti, non esistono confini netti e, soprattutto, se vi si sta immersi è quasi impossibile riconoscerne l’articolazione. Ma, se si guarda da fuori (e da lontano), ecco che una morfologia e un confine emerge chiaramente. Allo stesso modo, nella galassia DH possiamo individuare:

- il *nucleo*: un ambito quasi-disciplinare autonomo, che trascende le discipline tradizionali, focalizzato su questioni teoriche e metodologiche generali e trans-disciplinari e su aspetti infrastrutturali sia in senso epistemologico sia in senso pragmatico e istituzionale;

- il *disco*: il luogo in cui si collocano, come i bracci di spirale, le versioni informatiche “forti” delle discipline tradizionali - che spesso co-operano in modo interdisciplinare - (linguistica computazionale, filologia digitale, storia digitale, studi letterari computazionali, etc.), in cui i metodi computazionali e gli strumenti digitali hanno un riconosciuto ruolo metodologico e propongono questioni di ricerca e spiegazioni/interpretazioni innovative;
- il *bordo esterno*: lo spazio di confine in cui si dispongono le discipline tradizionali che accettano e includono i metodi e i risultati della ricerca informatizzata, e in cui si entra in contatto con altre galassie disciplinari come l’informatica, le *Library and Information Sciences*, i *media studies*, le scienze della comunicazione, le scienze cognitive.

Come tutte le metafore anche questa va presa per quello che è: un modo di rendere visibile e dicibile una complessità multidimensionale. Quello che mi preme in questa sede, tuttavia, è definire meglio il contenuto del nucleo, poiché credo che da questa definizione dipenda la possibilità di rivendicare uno spazio disciplinare autonomo. In estrema sintesi i contenuti caratterizzanti di tale nucleo, a mio parere, sono i seguenti:

- lo studio dei metodi computazionali (qualitativo/simbolici e quantitativi), e della loro relazione con le teorie e i concetti fondanti delle discipline umanistiche, sia latamente intese, sia prese nelle loro rispettive specificità, ma a un elevato livello di astrazione;
- la ricerca su modelli formali e sulle strutture di dati complesse richieste dalla formalizzazione nei domini umanistici;
- lo studio critico (in senso kantiano) delle interconnessioni tra metodi, modelli e problemi della ricerca umanistica, sulle loro condizioni di applicabilità, sui criteri di validazione, sui presupposti teorici ed epistemologici che li validano;
- lo studio, la creazione e l’innovazione delle infrastrutture generali che rendono possibile la ricerca umanistica digitale e computazionale, inclusa la produzione e gestione delle risorse digitali culturali, e che ne garantiscono una effettiva fruibilità, disseminazione e validazione scientifica e sociale.

Penso che su questa base si possano indentificare i contorni di un ambito disciplinare vero e proprio, la cui denominazione potrebbe essere “Metodologia computazionale e digitale per le scienze umane”, oppure, ritornando alla nostra illustre tradizione “Informatica Umanistica”. Si tratta di una definizione ancora molto generale e che necessita di ulteriori approfondimenti e giustificazioni, nonché di una esauriente esemplificazione in ricerche ed elaborazioni individuabili nel panorama degli studi digitali. Non intendo, infatti, proporre una ipostatica visione astratta che non tenga conto di ciò che accade nella reale dinamica della ricerca e della didattica, nella nostra realtà effettuale insomma.

## **CONCLUSIONI: UN POSSIBILE ESITO PER LA COLLOCAZIONE DISCIPLINARE IN ITALIA**

Come detto in apertura, questa ennesima riflessione meta-disciplinare non ha una finalità puramente teorica. Lo scopo che perseguo è verificare se sussistano le condizioni per la costituzione di ambito disciplinare autonomo anche in senso istituzionale, riconosciuto in sede istituzionale nel panorama accademica italiano (il problema per gli enti di ricerca è forse meno impellente, ma il discorso vela anche per quell’ambito), e che dunque possa essere un programma politico per la nostra associazione.

Per ottenere questo risultato la prima, fondamentale domanda è: è possibile individuare uno specifico insieme di attività di ricerca sufficientemente differenziato dalla tradizione e



coerente che possa ambire allo status disciplinare? Ebbene, quanto ho scritto ambisce a fornire una risposta positiva a questo quesito.

Ma questa è ovviamente solo una condizione necessaria ma affatto sufficiente allo scopo. Il resto, la questione difficile, è costituito dalla costruzione delle condizioni politiche e istituzionali di contorno, dai problemi giuridici e regolamentari, dalla costruzione di alleanze e dalla individuazione di compagni di strada. Da questo punto di vista la situazione attuale non è certo rosea, mi si possono forse intravedere degli spiragli. Il Ministero dell'Università lo scorso anno ha affidato al CUN una inchiesta esplorativa per la ridefinizione dei campi del sapere: non sono in grado oggi di dire se il mutato quadro politico consentirà di proseguire verso una vera fase di innovazione, e a essere sincero non poche sono le resistenze interne all'accademia stessa. Ma il lavoro elaborato fin qui apre interessanti prospettive per affrontare alcuni degli ostacoli regolamentari e istituzionali maggiori che si stagliano di fronte alla costituzione di un gruppo disciplinare autonomo, senza eccessivi velleitarismi (mi riferisco alla ventilata possibilità di costituire settori inter-area, che per il nostro campo di studi sarebbe essenziale).

Assumendo dunque un atteggiamento ottimistico riguardo l'agibilità istituzionale del progetto, quali potrebbero essere gli alleati e i compagni di strada? Penso che essi vadano individuati tra i settori a orientamento spiccatamente teorico e metodologico nelle discipline di area letteraria linguistica e storica, in una parte del settore archivistico e biblioteconomico e in alcuni esponenti intellettualmente "coraggiosi" provenienti dalle aree informatiche e filosofiche. In questo modo il nucleo di un settore disciplinare si potrebbe formare, in attesa che le nuove generazioni di studiosi contribuiscano a dargli un profilo scientifico più autonomo. Il tentativo di fondare una disciplina per gli studi umanistico digitali è stato già fatto in passato da studiosi ben più prestigiosi e autorevoli di noi, *in primis* Tito Orlandi, senza successo. Credo che sia il tempo di riprendere in mano questa battaglia, per rendere omaggio a questa importante eredità intellettuale.

## BIBLIOGRAFIA

- [1] Allington, Daniel, Sarah Brouillette, e David Golumbia. 2016. «Neoliberal Tools (and Archives): A Political History of Digital Humanities». *Los Angeles Review of Books (LARB)*, n. 1.5.2016. <https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>.
- [2] Berry, David M., a c. di. 2012. *Understanding Digital Humanities*. Palgrave Macmillan. [http://www.palgrave.com/page/detail/?sfl=id\\_product&st1=493310](http://www.palgrave.com/page/detail/?sfl=id_product&st1=493310).
- [3] Ciotti, Fabio. 2018. «Dall'Informatica umanistica alle Digital Humanities. Per una storia concettuale delle DH in Italia.» In *Digital Humanities 2018, DH 2018, Book of Abstracts, El Colegio de México, UNAM, and RedHD, Mexico City, Mexico, June 26-29, 2018*, 174–77. <https://dh2018.adho.org/en/dallinformatica-umanistica-alle-digital-humanities-per-una-storia-concettuale-delle-dh-in-italia/>.
- [4] Fish, Stanley. 2012. «The Digital Humanities and the Transcending of Mortality». *The New York Times «Opinionator»*, 9 gennaio 2012. <http://opinionator.blogs.nytimes.com/2012/01/09/the-digital-humanities-and-the-transcending-of-mortality/?hp>.
- [5] Gold, Matthew K., a c. di. 2012. *Debates in the Digital Humanities*. Univ Of Minnesota Press. <http://dhdebates.gc.cuny.edu/>.
- [6] Gold, Matthew K., e Lauren F. Klein, a c. di. 2016. *Debates in the Digital Humanities: 2016*. Minneapolis London: University of Minnesota Press.
- [7] Kirschenbaum, Matthew. 2010. «What Is Digital Humanities and What's It Doing in English Departments?» *ADE Bulletin* 150: 55–61.

- [8] Marche, Stephen. 2012. «Literature Is Not Data: Against Digital Humanities». *Los Angeles Review of Books*, 28 ottobre 2012. <http://lareviewofbooks.org/essay/literature-is-not-data-against-digital-humanities#>.
- [9] Nowviskie, Bethany. 2016. «On the Origin of “Hack” and “Yack”». In *Debates in the Digital Humanities 2016*.
- [9] McCarty, Willard, e Harold Short. 2002. «Mapping the Field». EADH. 2002. <http://www.eadh.org/mapping-field>.
- [10] Ramsay, Stephen. 2013. «Who’s In and Who’s Out». In *Defining Digital Humanities - A Reader*, a cura di Melissa Terras, Julianne Nyhan, e Edward Vanhoutte, 239–41. Williston: Ashgate. <http://www.ashgate.com/isbn/9781409469636>.
- [11] Ryan, Cordell. 2016. «How Not to Teach Digital Humanities». In *Debates in the Digital Humanities: 2016*, a cura di Matthew K. Gold e Lauren F. Klein. Minneapolis London: University of Minnesota Press. <http://dhdebates.gc.cuny.edu/debates/text/87>.
- [12] Siemens, Ray. 2016. «Communities of practice, the methodological commons, and digital self-determination in the Humanities». In *Doing Digital Humanities: Practice, Training, Research*, a cura di Constance Crompton, Richard J. Lane, e Ray Siemens, XXI–XXXIII. Taylor & Francis. <https://books.google.it/books?id=l-FRDQAAQBAJ>.
- [13] Siemens, Ray, John Unsworth, e Susan Schreibman, a c. di. 2004. *A Companion to Digital Humanities*. Hardcover. Blackwell Companions to Literature and Culture. Oxford: Blackwell Publishing Professional. <http://www.digitalhumanities.org/companion/>.
- [14] Svensson, Patrik. 2016. *Big Digital Humanities: Imagining a Meeting Place for the Humanities and the Digital*. University of Michigan Press. <https://doi.org/10.3998/dh.13607060.0001.001>.
- [15] Terras, Melissa, Julianne Nyhan, e Edward Vanhoutte, a c. di. 2013. *Defining Digital Humanities - A Reader*. Williston: Ashgate. <http://www.ashgate.com/isbn/9781409469636>.
- [16] Tomasin, Lorenzo. 2017. *L'impronta digitale: cultura umanistica e tecnologia*. 1a edizione. Sfere extra. Roma: Carocci editore.
- [17] Unsworth, John. 2000. «Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?» In . King’s College London. <http://people.brandeis.edu/~unsworth/Kings.5-00/primitives.html>.
- [18] ———. 2002. «What Is Humanities Computing and What Is Not?» *Jahrbuch Für Computerphilologie* 4 (2002), n. 4: 71–83.

# Per un archivio digitale dell'Università Castrense di San Giorgio di Nogaro: saperi, pratiche, immagini durante la Grande guerra

Dario De Santis

Dipartimento di studi umanistici e del patrimonio culturale, Università di Udine, dario.desantis@uniud.it

## ABSTRACT

Il presente contributo intende illustrare le fasi della realizzazione dell'Archivio digitale dell'Università Castrense (1916-1917) di San Giorgio di Nogaro (UD) e proporre alcune riflessioni emerse nel corso dei lavori. La documentazione relativa alla storia della Castrense, un particolare campus scientifico-universitario, ente di formazione e di ricerca sorto nel corso della Grande Guerra sulla linea del fuoco e distrutto a seguito della disfatta di Caporetto, si trova oggi disseminata in numerosi archivi di enti e di privati. La costituzione di un archivio digitale "tematico" e la creazione di un inventario virtuale analitico hanno messo in evidenza potenzialità e criticità di questo innovativo strumento di ricerca, didattica e divulgazione: un luogo virtuale di consultazione costantemente aggiornabile che traccia e rappresenta l'avanzamento delle ricerche e consente una fruizione efficace di documenti eterogenei. Emergono dal lavoro svolto nuove metodologie di ricerca, di narrazione e di valorizzazione delle fonti, nonché elementi di riflessione teorici e spunti applicativi per la determinazione di criteri di lavoro e fruizione più efficaci.

## PAROLE CHIAVE:

Archivi digitali, Divulgazione, Public History, Storia della medicina, Prima guerra mondiale.

## INTRODUZIONE

L'Università Castrense (o Scuola medica da campo) fu un campus universitario-militare di Medicina e Chirurgia sorto nel gennaio 1916 proprio a San Giorgio di Nogaro, che combinava formazione teorico-scientifica e applicazione pratica. Un'esperienza accademica correlata alle vicende belliche, con ricadute su ricerca scientifica, territorio italiano e concezione della guerra che permise a più di mille studenti universitari di terminare il percorso universitario al fronte e di prestare servizio in prima linea, forti di una preparazione medica specifica.

Nel corso del 2018 è stato avviato dal Dipartimento di studi umanistici e del patrimonio culturale (DIUM) dell'Università di Udine un progetto volto alla realizzazione dell'archivio digitale della Castrense, nell'ambito di un più ampio programma di ricerca, coordinato dal Prof. Francesco Pitassio, che si pone come obiettivo complessivo lo studio transdisciplinare del trauma nell'esperienza della Prima Guerra Mondiale. Grazie ad un finanziamento di Regione autonoma Friuli Venezia Giulia per progetti riguardanti la realizzazione di studi e ricerche storiche di base concernenti la prima guerra mondiale e del Comune di San Giorgio di Nogaro (UD) è stato innanzitutto digitalizzato e pubblicato online, su una piattaforma appositamente predisposta, il fondo documentario recentemente acquisito dalla Biblioteca comunale "Villa Dora" di San Giorgio di Nogaro.

Il progetto di digitalizzazione dell'archivio, di natura interdisciplinare, coinvolge ricercatori e professori dell'Università di Udine che appartengono a diversi settori scientifici del Dipartimento di Studi umanistici e del Patrimonio culturale (DIUM) e del Dipartimento di Scienze economiche e statistiche (DIES), e mira ai seguenti obiettivi:

1) Condizionamento, riordino e descrizione dell'archivio storico dell'Università Castrense;

- 2) Acquisizione digitale dei materiali, compresi quelli fotografici;
- 3) Costituzione dell'archivio digitale dell'Università Castrense;
- 4) Valorizzazione dell'archivio mediante pubblicazione e diffusione all'interno di un sito web appositamente predisposto.

L'archivio digitale sarà presentato ufficialmente nella primavera del 2019, in occasione del convegno internazionale dedicato al concetto di trauma nel corso della Grande Guerra, che si terrà a Udine (presso il DIUM) e a San Giorgio di Nogaro. Successivamente l'archivio verrà pubblicato online come sezione di primo livello del sito dell'Università Castrense, già creato dalla biblioteca comunale Villa Dora: [www.universitacastrense.eu](http://www.universitacastrense.eu)

### ***Archiui: una nuova piattaforma***

Per la pubblicazione online delle carte è stato utilizzato il sistema *Archiui*, realizzato dalla software house Promemoria s.r.l. di Torino. Tale database è composto dal software di descrizione *opensource* CollectiveAccess – nato inizialmente negli Stati Uniti per gestire il patrimonio museale e in seguito sviluppato anche per le altre tipologie di beni culturali – integrato con il content management system WordPress, molto diffuso anche nel mondo editoriale e particolarmente apprezzato per la facilità d'uso e la buona indicizzazione dei contenuti sui motori di ricerca. In sostanza per la visualizzazione on-line di tutti i contenuti (front-end), viene utilizzato WordPress, grazie ad un plugin – sviluppato dalla stessa azienda – che permette di integrare all'interno della medesima interfaccia web contenuti di tipo editoriale e contenuti che vengono richiamati in modo dinamico dal database di CollectiveAccess.

*Archiui* è oggi una piattaforma stabile, accessibile e ampiamente personalizzabile. CollectiveAccess rispetta gli standard internazionali di descrizione archivistica e biblioteconomica (ISAD, ISAAR, DublinCore, DACS, MARC, PBCore e altri) e consente una gestione unificata di diverse banche dati, tracciando quelle preesistenti in caso di migrazione da altri applicativi (Archimista, Guarini Archivi, GEA, Arianna, MS Access, MS Excel), garantendo l'interoperabilità e la condivisione delle risorse con altri sistemi informativi. Forte di una community internazionale ampia, consente di esportare i dati codificati XML-EAD e XML-EAC, mentre l'ambiente software è indipendente dalla piattaforma di installazione, che può essere proprietaria (Microsoft Windows Server 2003, Server 2008, Windows XP e Windows 7, Solaris 9+ o Mac OS X 10.5+) oppure open source (Linux, nelle diverse distribuzioni). Cfr. <http://docplayer.it/959212-Collectiveaccess-l-open-source-al-servizio-degli-archivi-storici.html>.

### **L'ARCHIVIO DIGITALE DELL'UNIVERSITÀ CASTRENSE: UN PERCORSO “TEMATICO”**

Riperkorrendo tutte le fasi del progetto, la presentazione intende porre l'accento su alcune peculiarità dell'archivio digitale dell'Università Castrense che si inseriscono nella più ampia riflessione sulle nuove prospettive di valorizzazione e divulgazione offerte dall'universo digitale.

Il nucleo documentale oggetto del riordino e della digitalizzazione consiste in circa duemila carte rinvenute a Sarzana in due valigie conservate all'interno dell'abitazione di Giuseppe Tusini, professore di clinica chirurgica presso l'Università di Modena, fondatore e direttore dell'Università Castrense nel 1916, donate dagli eredi alla biblioteca civica “Villa Dora” di San Giorgio di Nogaro. Tali documenti costituiscono una fonte di primaria importanza per lo studio delle vicende che portarono alla nascita di questo atipico centro di formazione e di ricerca e consentono di approfondire la disamina dell'attività didattica e scientifica svolta all'interno del campus nel corso dei suoi due anni di attività.

Nell'archivio troviamo: l'elenco dei corsi e degli iscritti, i verbali degli esami, l'attività delle cliniche universitarie ubicate negli ospedali da campo, le relazioni redatte dal direttore e dal corpo docenti, i verbali degli esami e delle discussioni di laurea, la corrispondenza del direttore, le bozze delle pubblicazioni scientifiche emerse dall'attività di ricerca, ma anche la documentazione raccolta dallo stesso Tusini (verbali delle sedute parlamentari, appunti, articoli di giornale, corrispondenza) per ricostruire i passaggi istituzionali che portarono all'approvazione parlamentare e al disegno di legge.

In una prima fase del lavoro si è provveduto a completare il riordino fisico del fondo, già avviato dal personale della biblioteca, e a terminare il suo condizionamento. Successivamente è stato possibile digitalizzare i documenti e provvedere alla realizzazione di un inventario analitico.

Contestualmente è stata condotta su tutto il territorio nazionale una ricerca volta a individuare documenti e materiali fotografici relativi alla Castrense e ai suoi protagonisti. Tale scavo ha portato all'individuazione di numerose unità archivistiche sia in archivi di enti sia in archivi di privati, per un totale di circa 6.500 carte, che sono state acquisite in formato digitale ad alta definizione. Numerosi documenti sono stati rinvenuti ad esempio presso l'AUSME - Archivio documentale dello Stato maggio dell'Esercito (Fondi E7, B3, B1, F3), l'Archivio centrale dello Stato (Archivio del Ministero della pubblica istruzione) e l'Archivio Storico della Croce Rossa a Roma. Altri archivi, soprattutto di privati, come l'Archivio personale dei docenti Pietro Marogna (Sassari), Gherardo Forni (San Giovanni in Persiceto) e Gaetano Samperi (Catania), che contengono diari, corrispondenza e materiale fotografico relativo alla Castrense sono stati individuati e segnalati e potranno essere aggregati in futuro.

Considerata la natura istituzionale e giuridica straordinaria dell'Università Castrense – afferente in parte al Ministero della Guerra, in parte a quello dell'Istruzione, prima a statuto autonomo, poi succursale dell'Università di Padova, distrutta a seguito della disfatta di Caporetto – la ricerca delle fonti documentarie legate alla sua storia presenta notevoli difficoltà. Per questa ragione, con la creazione di un archivio digitale e a partire dal nucleo documentale recentemente rinvenuto, si è ritenuto opportuno pubblicare online tutte le fonti emerse dallo scavo.

L'archivio digitale della Castrense è dunque oggi costituito da un “superfondo” (per un totale di circa 8.000 carte digitalizzate) che, oltre alla documentazione conservata a Villa Dora, propone unità archivistiche provenienti da diversi fondi (6 attualmente, ma in costante aggiornamento), sparsi sul territorio nazionale. La dimensione virtuale infatti consente agli utenti di navigare e interagire con l'informazione in modi nuovi, mettendo in discussione gli attuali fondamenti epistemologici delle strutture documentarie che potrebbero essere ripensate incoraggiando modalità di studio e interpretazione delle collezioni che rispecchino anche il punto di vista e le aspettative del pubblico. All'interno di questo scenario, gli archivi digitali possono essere considerati terreno fertile per la sperimentazione di modelli innovativi di accesso e fruizione del patrimonio culturale: l'archivio digitale della Castrense presenta in questo senso una struttura “tematica” che consente agli studiosi di consultare la documentazione relativa a uno specifico percorso di ricerca. Tale struttura è oggi ancora relativamente poco diffusa: nel panorama digitale prevalgono infatti archivi online che riproducono la struttura fisica di uno o più fondi – è il caso ad esempio dell'Aspi - Archivio storico della psicologia italiana – o banche dati cumulative che raccolgono e presentano il lavoro di diversi gruppi di ricerca, come ad esempio il progetto *Cendari*. In questo caso invece si è deciso di adottare una “terza via”: oltre alla digitalizzazione e alla pubblicazione online del nucleo documentale principale, l'archivio storico dell'Università Castrense, sono stati digitalizzate partizioni, serie o semplici unità archivistiche appartenenti a fondi di istituzioni o di privati, legate all'attività istituzionale, militare, clinica o scientifica dell'Università Castrense o dei suoi protagonisti. Pur mantenendo la sua funzione originaria –

conservazione (digitale) e consultazione – l’archivio acquisisce così una funzione nuova e diviene un ambiente virtuale di lavoro dove una comunità sempre più ampia ritrova i documenti d’archivio relativi a uno specifico percorso di ricerca all’interno di una struttura archivistica che garantisce il mantenimento delle relazioni d’archivio originarie e consente di individuare facilmente l’ubicazione fisica della documentazione e la sua collocazione.

In questa veste l’archivio digitale diventa non solo il luogo di conservazione, consultazione e valorizzazione di un determinato fondo, ma traccia un percorso di indagine che può essere costantemente aggiornato e ampliato: uno strumento innovativo utile non solo agli studiosi ma anche all’attività didattica e alla *hands-on-practice*. Registra e mostra infatti i “movimenti” degli storici, le relazioni complesse e spesso casuali che legano gli avvenimenti e dunque i documenti e gli archivi; fornisce una rappresentazione dell’attività di ricerca, riportando, in un unico inventario virtuale e nella loro disposizione originaria, le fonti su cui si basano le diverse interpretazioni e conclusioni degli studiosi. In un ambiente in cui i vincoli fisici sono potenzialmente inesistenti, i tradizionali modelli e approcci usati dagli specialisti possono essere integrati con nuovi linguaggi e strategie di organizzazione e visualizzazione dei contenuti. È certo che l’approdo della documentazione digitale nei nostri computer sotto forma di risorse sia in grado di trasformare il modo di lavorare dello storico ma anche il suo rapporto con la scrittura e con i documenti. In questa veste, l’archivio stesso, oltre a strumento di indagine, diviene narrazione della ricerca svolta e permette a una comunità sempre più ampia di ripercorrere ed eventualmente proseguire il lavoro, consentendo l’utilizzo delle fonti per la didattica e favorendo nuove modalità di valorizzazione e divulgazione.

## OBIETTIVI

La presentazione intende dunque:

1. Illustrare lo svolgimento del progetto in tutte le sue fasi (riordino, censimento di documenti sul territorio, costituzione dell’archivio digitale “tematico”) per “raccontare” la creazione di un archivio digitale.
2. Mostrare l’archivio online e il suo funzionamento.
3. Alla luce delle considerazioni qui esposte riflettere su potenzialità/criticità di un ambiente digitale “tematico” e sul suo utilizzo in ambito didattico e divulgativo.
4. Con particolare attenzione alla *Public History*, illustrare come il passaggio concettuale tra il mondo reale e quello digitale, trasformando il modo di rapportarsi a fonti e documenti, offra innovativi strumenti di ricerca e narrazione.

## BIBLIOGRAFIA

- [1] Allegrezza, Stefano. *Le memorie per la conservazione a lungo termine dei documenti digitali: metodi e criteri per la valutazione e la selezione dei supporti di memorizzazione*. Macerata: Simple, 2013.
- [2] Allegrezza Stefano, Gorgolini, Luca (a cura di). *Gli archivi di persona nell'era digitale: il caso dell'archivio di Massimo Vannucci*. Bologna: Il Mulino, 2016.
- [3] Bollini, Letizia. “Lo spazio digitale del sapere: dalla forma archivio alle knowledge-base culturali. Aggregazioni, narrazioni e migrazioni”. In *Design & Cultural Heritage*, a cura di Raffaella Trocchianesi ed Eleonora Lupo, 3:53-67. Milano: Electa, 2013.
- [4] Bollini, Letizia. “The open heritage of digital archives. Preservation, sustainability and accessibility of historical documentation to map the field of Italian Mind Science”. In *Heritage 2016*, a cura di Rogério Amôêda, Sérgio Lira e Cristina Pinheiro, 773-782. Lisbona: Green Lines Institute, 2016.
- [5] Bollini, Letizia e Marco Borsotti. “Strategies of Commutation in Exhibition Design”. *The International Journal of Architectonic, Spatial, and Environmental Design* 10, num. 1 (2016): 13-21.
- [6] Burdick, Anne, Johanna Drucker, Peter Lunenfeld, Todd Presner, e Jeffrey Schnapp. *Digital humanities*. Cambridge, MA: MIT Press, 2012.
- [7] Bollini, Letizia, De Santis, Dario, Radice, Sara e Zocchi, Paola. “Le trame invisibili. Nuove modalità di esplorazione online dell’Archivio storico della psicologia italiana”. *Umanistica digitale*, num. 1, 2017: 59-84.

- [8] Cameron, Fiona. "Museum Collections, Documentation, and Shifting Knowledge Paradigms". In *Reinventing the Museum: The Evolving Conversation on the Paradigm Shift*, a cura di Gail Anderson, Seconda, 223-238. Plymouth: AltaMira Press, 2012.
- [9] De Santis, Dario. "L'archivio storico della psicologia italiana". *Rivista di storia della medicina* 20, num. 1-2 (2010): 161-163.
- [10] Guercio, Maria. *Archivistica informatica. I documenti in ambiente digitale*. Roma: Carocci, 2015.
- [11] Guercio, Maria. *Conservare il digitale. Principi, metodi e procedure per la conservazione a lungo termine di documenti digitali*. Roma-Bari: Laterza, 2013.
- [12] Nicks, John. "Curatorship in the Exhibition Planning Process". In *The Manual of Museum Exhibitions*, a cura di Barry Lord e Gail Dexter Lord, 345-72. Walnut Creek, CA: AltaMira Press, 2002.
- [13] Noiret Serge. "Informatica, storia e storiografia: la storia si fa digitale". *Memorie e ricerca*, num. 28, maggio-agosto 2008: 189-201.
- [14] Radice, Sara. "The open archive: toward a framework for the design of virtual exhibition of archival contents enabling an effective audience engagement". In *The Ecology of Culture: Community Engagement, Co-creation, Cross Fertilization*, 392-401. Brussels: ENCATC, 2015.
- [15] Salvadori, Enrica. "Digital (Public) History: la nuova strada di un'antica disciplina. *Rivista dell'Istituto di storia dell'Europa mediterranea*, numero 1/I n.s., dicembre 2017: 57-94.
- [16] Tosti Croce, Mauro e Maria Natalina Trivisano. "Thematic Portals: Tools for Research and Making the Archival Heritage Known". *Journal of Modern Italian Studies* 20, num. 5 (2015): 732-739.
- [17] Vitali, Stefano. *Passato Digitale. Le fonti dello storico nell'era del computer*. Milano: Bruno Mondadori, 2004
- [18] Zanni Rosiello, Isabella. "A proposito di web e del mestiere di storico". *Contemporanea*, VIII, num. 4, Ottobre 2005:743-755.
- [19] Walsh, Peter. "Rise and Fall of the Post-Photographic Museum: Technology and the Transformation of Art". In *Theorizing digital cultural heritage: a critical discourse*, a cura di Fiona Cameron e Sarah Kenderdine, 19-34. Cambridge, Mass: MIT Press, 2007.
- [20] Zocchi, Paola. "Gli archivi storici della psicologia italiana on-line". *Storia in Lombardia* 1 (2010): 150-51.
- [21] Zocchi, Paola. "Il progetto «Archivi storici della psicologia italiana». *Archivi* 2 (2014): 119-130.

# Funzione dell'edizione critica digitale nel contesto degli studi universitari. Questioni di scelte didattiche: studenti codificatori o studenti esecutori?

Luciano Longo, Alba Castello, Natalia Librizzi, Ugo La Mantia  
[longo@olomedia.it](mailto:longo@olomedia.it); [alba.castello@unipa.it](mailto:alba.castello@unipa.it);  
[natalia.librizzi@unipa.it](mailto:natalia.librizzi@unipa.it); [ugolamant@gmail.com](mailto:ugolamant@gmail.com);

## ABSTRACT

Il presente contributo si pone l'obiettivo di riflettere sulla funzione didattica di una edizione critica digitale nel contesto degli studi universitari. La riflessione prova a ragionare sulla possibilità di diversificare l'accesso degli studenti alla realizzazione di una edizione critica utilizzando la sintassi XML-TEI. Le vie individuate e sperimentabili sono sostanzialmente due: la prima, già ampiamente usata, si basa su una acquisizione diretta degli elementi di codifica da parte dello studente; la seconda, meno praticata, invece si basa sul fornire allo studente una struttura esecutiva su cui approntare e rappresentare l'edizione. Dunque, la presente ricerca parte dal problema capitale nel campo delle *digital humanities* che riguarda la rappresentazione e la visualizzazione del processo di scrittura utilizzando le linee guida della TEI da parte di uno studente universitario. A questo problema si associa una riflessione più ampia, e per certi versi impervia, sulla natura dello studioso che vuole realizzare una edizione critica digitale. In questa occasione si intende porre una questione non più procrastinabile: lo studente-studioso deve essere a tutti i costi un codificatore di testi o può essere un esecutore di strumenti già strutturati? In sintesi si tratta di ragionare su due modelli didattici diversi ma non contrapposti: uno basato sull'acquisizione del metodo e l'altro basato sull'uso dello strumento di visualizzazione.

## PAROLE CHIAVE

Testo; tool; metodo; strumenti

## PROPOSTA

La proposta di contributo parte da alcune acquisizioni teorico-metodologiche ormai consolidate. È importante partire dalla considerazione che l'applicazione dei linguaggi di codifica permette una diversa prospettiva nella caratterizzazione del testo inteso come «processo» e un passo in avanti nella visualizzazione anche comparativa dei testimoni. Il testo codificato non solo presenta un testo univoco, cioè il *text based*, obiettivo dell'ecdotica tradizionale, ma anche un ipotesto, cioè un testo che contiene in sé molti testi e molteplici situazioni testuali. Infatti, negli ultimi anni l'attenzione della ricerca si è concentrata fondamentalmente sulla *funzione* del testo e sulla sua natura *dinamica*; oggi non si riflette più sull'*unicità o esemplarità del testo tradito* ma sulla *naturale molteplicità* del soggetto testo. L'edizione digitale con la marcatura in XML-TEI e l'utilizzo di *tool* specifici non ha solo la funzione di rappresentare il documento originale, caratteristica già presente nelle edizioni standard, ma ha la funzione soprattutto di rappresentare il testo come «movimento di scrittura» e la scrittura come «atto di pensiero complesso». Ciò non significa però che la creazione di una struttura di *meta-testo* debba prevalere sulle esigenze del lettore; per mediare tra rappresentazione del testo e la sua fruizione è necessario *ripensare* l'edizione, e quindi la



sua strutturazione in *tag*, come etichettatura interpretativa del testo e del suo farsi «atto» e «attività», sperimentando una codifica XML-TEI anche su XSLT che produca *output* diversi per testare così la coerenza rappresentazionale del testo in direzione della sua fruizione.

In generale si può affermare che l'XML così come l'XSLT sono atti interpretativi,<sup>2</sup> in quanto la scelta di far visualizzare determinati elementi e attributi del testo codificato in un dato modo, corrisponde al modello di testo che l'editore vuole consegnare al lettore. Ma è più giusto, prendendo in prestito le parole di Pierazzo, definire l'XSLT «prove a concept», cioè una possibile rappresentazione della visualizzazione di fenomeni e di strutture testuali. Infatti, uno dei problemi più critici, nella strutturazione di una edizione digitale, è proprio la creazione dell'interfaccia utente. Ogni *tool*, che determina la visualizzazione della rappresentazione del testo marcato per l'utente, impone regole e misure che non sempre sono in linea con la base epistemologica definita dall'editore, o più semplicemente, con il modello di codifica adottato.

A tal proposito, nel corso della ricerca e della sperimentazione del modello teorico, che ha portato alla creazione delle esemplificazioni qui proposte, si è provato a visualizzare la codifica standardizzata per le diverse tipologie testuali su software diversi. Il risultato non sempre ha fatto emergere le caratteristiche del testo, ma soprattutto imponeva una ricodifica del documento e dei suoi fenomeni. È questo uno dei problemi più critici, e in parte anche il più scoraggiante, nel realizzare una edizione critica digitale; la codifica di un testo non sempre è esportabile su diversi *tool*, rendendo il lavoro del codificatore-interprete davvero difficoltoso e in alcuni casi privo di risultati meritevoli, tranne che lo studioso non si trasformi antropologicamente ancor più in un programmatore informatico. Il punto non sta tanto nella trasformazione del codificatore-interprete, quanto nella difficoltà, anche in termini temporali, di apprendere linguaggi e strutture che per formazione e competenze appartengono ad altri ambiti disciplinari. Il rischio maggiore di una tale trasformazione potrebbe determinare uno spostamento dell'attenzione del codificatore-interprete dalla filologia ad altro.

Questa situazione dello stato della ricerca deve necessariamente far riflettere su quali scelte e indirizzi bisogna intraprendere nel contesto degli studi universitari per la creazione di una edizione critica digitale. Sintetizzando, le vie percorribili sono due: la prima è quella di fornire e far acquisire agli studenti gli elementi della sintassi XML-TEI e provare a far codificare le diverse tipologie testuali; la seconda è quella di formalizzare, tramite un processo di standardizzazione, tipologie testuali e far lavorare gli studenti non tanto sull'azione di codifica quanto su quella della visualizzazione e rappresentazione del testo critico. Questa seconda opzione prevede la creazione di fogli XML già codificati e formalizzati sui quali gli studenti dovrebbero apportare o inserire, secondo criteri epistemologici ben definiti, il sistema variantistico del testo di cui vogliono realizzare l'edizione critica. Questa linea di formazione sposterebbe la scelta didattica da uno studente codificatore a uno studente esecutore.

Il contributo prova a presentare l'itinerario progettuale su cui si sta lavorando all'Università di Palermo basato su due linee: da un lato, quella di codificare diverse tipologie di testo secondo la sintassi XML-TEI; dall'altro, di avviare una sperimentazione di edizioni critiche digitali basate su modelli di codifica già standardizzati e verificati su diversi *tool*. La prima linea di ricerca e di sperimentazione, cioè quella relativa alle tipologie testuali, si basa su quattro codifiche testuali; le prime due riguardano la codifica di testi poetici, le seconde testi in prosa. La prima tipologia testuale riguarda testi poetici definiti in un progetto d'autore e testimoniati sia da manoscritti che da dattiloscritti. I componimenti presentati esibiscono diverse evidenze scritte, dalla cancellatura e riscrittura di singole lezioni fino all'espunzione di interi versi o segmenti strofici; dall'aggiunta nella linea di scrittura a quella nell'interlinea o nei margini

---

<sup>2</sup> Ciotti, F., *Breve introduzione alla Text Encoding Initiative*, in *Biblioteca Italiana*.

fino alla ri-dislocazione di versi di varia misura da una posizione all'interno dello stesso componimento a testi che assumono nuova forma. La presente tipologia testuale prende in considerazione il caso delle liriche di Lucio Piccolo (1901-1969) già edite nella raccolta intitolata *Plumelia* (1967). Di tali componimenti verrà fornita una codifica di base per formalizzare la struttura partitica in strofe e versi e i diversi livelli di intervento autoriale (espunzione e aggiunte). Dunque, verranno esplicitate tutte le sequenze di codifica che, a partire dall'analisi di testimoni dattiloscritti e manoscritti, hanno permesso di rappresentare i movimenti alla base del processo di costruzione del testo fino alla sua stabilizzazione nella forma attestata nella prima edizione. Per questa tipologia testuale si esplicherà la codifica del modulo *Apparatus* secondo le impostazioni *location-referenced method*; questo sistema di marcatura permette di codificare e istituire, attraverso l'attributo di localizzazione @loc e i relativi valori, rappresentanti degli 'identificativi di relazione', delle correlazioni tra porzioni di testo che nelle varie redazioni si trovano in luoghi differenti ma tra le quali, secondo criteri formali, contenutistici e/o tematici, è presente una chiara corrispondenza.

La seconda tipologia testuale riguarda testi poetici inediti non inseriti in un progetto d'autore e testimoniati sia da manoscritti che da dattiloscritti. I componimenti presentati esibiscono diverse evidenze scritte, dalla cancellatura e riscrittura di singole lezioni fino all'espunzione di interi versi o segmenti strofici; dall'aggiunta nella linea di scrittura a quella nell'interlinea o nei margini. L'elemento caratterizzante di questa tipologia di componimenti è la presenza di situazioni scritte associabili all'equivalenza variantiva o all'adiazione. La presente tipologia testuale prende in considerazione il caso delle liriche inedite di Bartolo Cattafi (1922-1979) databili negli anni '60 del Novecento. Di queste poesie si esplicheranno, oltre i dati di codifica espressi con l'esemplificazione di testi de *L'osso, l'anima* (1964), anche i tag di codifica utilizzabili per testi in cui non è espressa una volontà definitiva d'autore. Per questa tipologia testuale si esplicherà la codifica del modulo *Apparatus* secondo le impostazioni *parallel-segmentation*. Tale modalità è finalizzata a esemplificare il testo come soggetto dinamico e fluido.

La terza tipologia riguarda testi narrativi incompiuti e inediti che presentano una situazione di tradizione a testimone unico. I testi esibiscono diverse evidenze scritte, dalla cancellatura e riscrittura di singole lezioni fino all'espunzione di intere sezioni. La presente tipologia testuale prende in considerazione il progetto narrativo *Enea e Anchise* di Elio Vittorini, scritto con ogni probabilità tra il 1947 e il 1950. Di tale testo verrà fornita una codifica di base per formalizzare la struttura in prosa e i diversi livelli di intervento autoriale (espunzione e aggiunte). Verranno esplicitate tutte le sequenze di codifica che, a partire dall'analisi del testimone manoscritto, ha permesso di rappresentare i movimenti alla base del processo di costruzione del testo fino a una possibile situazione di approssimazione alla stabilità narrativa. Per questa tipologia testuale si propone un tag set esemplificativo per una edizione documentaria attraverso l'uso anche di un modulo specifico, *zone*, contemplato nella sintassi XML-TEI.

La quarta tipologia riguarda testi in prosa con una storia redazionale molto complessa, in cui vi sono almeno due testimoni manoscritti e più di tre redazioni a stampa. Questa tipologia ha la finalità di presentare una ipotesi di codifica di redazioni manoscritte che presentano sia una straordinaria molteplicità di interventi correttivi immediati a cui si sovrappongono interventi tardivi; sia varianti sovrapposte di tipo instaurativo, sostitutivo e ri-propositivo; sia una diversità di inchiostrazione; sia una struttura compositiva dei testimoni autografi, che non è quasi mai speculari alle realizzazioni a stampa. La presente tipologia testuale prende in considerazione il caso de *I Viceré* di Federico De Roberto. Infatti, la storia redazionale di questo romanzo è caratterizzata da cinque testimoni. La prima testimonianza è tramandata da un manoscritto datato 1892; una seconda da un manoscritto datato 1893-94. Vi sono poi tre edizioni a stampa: la I<sup>a</sup> è del 1894 ad opera dell'editore Vecchi Galli di Catania; la II<sup>a</sup> rivista è

del 1920, pubblicata a Milano dai F.lli Treves, e la III<sup>a</sup> è del 1935 (F.lli Treves, Milano, 1935) in cui l'autore apporta soprattutto variazioni di forma.

In conclusione, la presentazione delle quattro esemplificazioni ha la funzione di mostrare quattro possibili modelli di codifica già strutturati e utilizzabili da studenti senza un loro coinvolgimento diretto nella realizzazione, prima del modello teorico di codifica, e poi nella definizione del set tag da utilizzare. L'esperimento che si vuole avanzare prevede l'acquisizione delle strutture ecdotiche dei modelli esposti e l'utilizzo di queste per la realizzazioni di edizioni critiche digitale come ponte per intraprendere studi ermeneutici ed interpretativi.

## BIBLIOGRAFIA

- [1] Amoroso, G., 1988. *Lucio Piccolo. Figura d'enigma*, All'insegna del pesce d'oro, Milano.
- [2] Cattafi, B., 2001. *Poesie*, a cura di Giovanni Raboni e Vincenzo Leotta, Mondadori, Milano.
- [3] Burnard, L., Bauman, S., 2014. TEI P5: Guidelines for Electronic Text Encoding and Interchange, Text Encoding Initiative Consortium, Charlottesville, Virginia (SIGLA: GP5).
- [4] Burnard, L., Sperberg-McQueen, C. M., (trad. a cura di Fabio Ciotti), 2014. *TEI Lite: introduzione alla codifica dei testi*. ([http://www.tei-c.org/Vault/P4/Lite/teiu5\\_it.html](http://www.tei-c.org/Vault/P4/Lite/teiu5_it.html))
- [5] Burnard, L., 2014. *What is the Text Encoding Initiative? How to Add Intelligent Markup to Digital Resources.*, OpenEdition Press, Marseille.
- [6] Covi, R., 1998. *Il lungo viaggio di Vittorini. Una biografia critica*, Marsilio, Venezia.
- [7] D'Iorio, P., Ferrand, N., 1998. *Genesi, critica, edizione*, Scuola Normale Superiore, Pisa.
- [8] De Biasi, P.M., 1996. "Edition horizontale, édition verticale. Pour une typologie des éditions génétiques (le domaine français 1980-1995)", in *Éditer des manuscrits. Archives, complétude, lisibilité*, (a c. di Benrekassa, G., Neefs, J., Didier, B.), PUV, Vincennes, pp. 159-193.
- [9] De Biasi, P.M., 1996. What is a literary draft? Toward a functional typology of genetic documentation, in *Yale French Studies*, 89, pp. 26-58.
- [10] De Biasi, P.M., 2004. "Toward a Science of Literature: Manuscript Analysis and the Genesis of the Work", in *Genetic Criticism: Texts and Avant-Textes*, (a c. di Deppman, J., Ferrer, D., Groden, M.), University of Pennsylvania Press, Philadelphia, pp. 26-68.
- [11] Di Grado, A., 1980. *Il silenzio delle Madri. Vittorini da Conversazione in Sicilia al Sempione*, Edizioni del Prisma, Catania.
- [12] Fiormonte, D., 2003. *Scrittura e filologia nell'era digitale*, Bollati Boringhieri, Torino.
- [13] Fiormonte, D., 1995. "Varianti elettroniche." in *Italiano & oltre 2*.
- [14] Gigliozzi, G., 1987. *Studi di codifica e trattamento automatico di testi*, Bulzoni, Roma.
- [15] Grésillon, A., 1988. *De la genèse du texte littéraire*, Du Lérot, Parigi.
- [16] Isella, D., 1987. *Le carte mescolate. Esperienze di filologia d'autore*, Liviana, Padova.
- [17] Italia, P., 2016. *Il metodo Leopardi. Varianti e stile nella formazione delle Canzoni*, Carocci, Roma.
- [18] Lo Dico, O., 1984. *Elio Vittorini*, Edizioni Salvatore Sciascia, Caltanissetta-Roma.
- [19] Maccari, P., 2003. *Spalle al muro. La poesia di Bartolo Cattafi*, Società Editrice Fiorentina, Firenze.
- [20] McGann, J., 2006. *La letteratura dopo il www. Il testo letterario nell'era digitale*, Bonomia University Press, Bologna.
- [21] McGann, J., 2004. "Marking texts of many dimensions", in *A Companion to Digital Humanities*, (a c. di Schreibman S. - Siemens-Unsworth R.), Blackwell, Oxford, pp. 198-217.
- [22] Mordenti, R., 2001. *Informatica e critica dei testi*, Bulzoni, Roma.
- [23] Mordenti, R., 2013. *L'altra critica. La nuova critica della letteratura tra studi culturali, didattica e informatica*, Editori Riuniti university press, Roma.

- [24] Numerico, T., Vespignani, A., 2003. *Informatica per le scienze umanistiche*, Il Mulino, Bologna.
- [25] Numerico, T., Fiormonte, D., Tomasi, F., 2010. *L'umanista digitale*, Il Mulino, Bologna.
- [26] Panicali, A., 1994. *Elio Vittorini. La narrativa, la saggistica, le traduzioni, le riviste, l'attività editoriale*, Mursia, Milano.
- [27] Perrone, D., 2013. "Dalle "Città del mondo" al Manoscritto di Populonia. L'ultimo Vittorini", in *Il Giannone*, anno XI, n. 22, pp. 199-216.
- [28] Pierazzo, E., 2015. *Digital Scholarly Editing*, ASHGATE.
- [29] Pierazzo, E., 2005. *La codifica dei testi*, Carocci Editore, Roma.
- [30] Pierazzo, E., 2011. "A Rationale of Digital Documentary Editions." in *Literary e Linguistic Computing*, 26/4.
- [31] Pierazzo, E., 2014. "Of Time and Space: a New Framework for Digital Editions of Draft Manuscripts." in *Variants*.
- [32] Prandi, S., 2007. *Da un intervallo al buio. L'esperienza poetica di Bartolo Cattafi*, Manni, San Cesario di Lecce.
- [33] Rosselli Del Turco, R., 2015. "Edition Visualization Technology: A Simple Tool to Visualize TEI-Based Digital Editions", in *Journal of the Text Encoding Initiative Issue*, 8. (<http://jtei.revues.org/1077>; DOI: 10.4000/jtei.1077).
- [34] Sciascia, L., 1991. *Le «soledades» di Lucio Piccolo*, in *La corda pazza. Scrittori e cose della Sicilia*, Adelphi, Milano.
- [35] Sinatra, M.E., Vitali-Rosati, M., 2014. *Pratiques de l'édition numérique*, Les presses de l'Université de Montréal, Montréal.
- [36] Spadini, E., Barabucci, G., Turska, M., 2017. "Data vs. Presentation. What is the core of a Scholarly Digital Edition?", in *Advances in Digital Scholarly Editing*, Sidestone Press, Leiden, pp. 37-46.
- [37] Tedesco, N. (a cura di), 2000. *Pensiero e cultura letteraria dell'Ottocento e del Novecento in La Storia della Sicilia*, Editalia, Roma.
- [38] Tedesco, N. (a cura di), 1990. *Lucio Piccolo. La figura e l'opera*, Pungitopo, Marina di Patti.
- [39] Tedesco, N. (a cura di), 1999. *Lucio Piccolo Giuseppe Tomasi. Le ragioni della poesia, le ragioni della prosa*, Flaccovio, Palermo.
- [40] Tedesco, N., 1995. *Poeti siciliani del Novecento*, Flaccovio Editore, Palermo.
- [41] Tedesco, N., 2003. *Lucio Piccolo. Cultura della crisi e dormiveglia mediterraneo*, Sciascia editore, Caltanissetta-Roma.
- [42] Tomasi F., 2008. *Metodologie informatiche e discipline umanistiche*, Carocci, Roma.

# The Digital Repository Service of the National Documentation Centre in Greece: a model for Digital Humanities data management and representation

Katerina Bartzi<sup>1</sup>, Nikos Vasilogamvrakis<sup>2</sup>, Elena Lagoudi<sup>3</sup>, Despina Hardouveli<sup>4</sup>, Evi Sachini<sup>5</sup>

<sup>1</sup> National Documentation Centre / National Hellenic Research Foundation, Greece – bartzi@ekt.gr

<sup>2</sup> National Documentation Centre / National Hellenic Research Foundation, Greece – nikvasil@ekt.gr

<sup>3</sup> National Documentation Centre / National Hellenic Research Foundation, Greece – elena.lagoudi@ekt.gr

<sup>4</sup> National Documentation Centre / National Hellenic Research Foundation, Greece – dxardo@ekt.gr

<sup>5</sup> National Documentation Centre / National Hellenic Research Foundation, Greece – esachin@ekt.gr

## ABSTRACT

Digital Humanities lie at the crossroads between technology and humanistic research. Contemporary technological advances significantly open up new avenues for humanities study and new ways to utilize the output of humanities endeavors for the communities that benefit from it, be it communities of practice, the education system, citizen scientists and society overall.

The following article describes the development cycle of a new cloud service the Greek National Documentation Centre (EKT) offers to the domestic cultural and science organisations. The National Documentation Centre (EKT) is one of the most important digital content stakeholders in Greece, facilitating Humanities research through robust e-infrastructures and widening access to digital collections.

The article describes the Digital Repository Service development from planning to delivery, describes the various components of the service and addresses the issues of engaging with the community of users to address their documentation needs and preferences. As part of designing the service and workflows, a survey was conducted in the GLAM sector organisations in Greece, in order to assess user needs and understanding of new and emerging technologies, as well as discern possible volume of digital content. Extensive literature and similar trends review took place and it was funneled into designing the documentation strategy used for the Repositories. The accompanying suite of services was developed (eLearning, eKnowledgeBase) and the project rolled out initially with some pilots, then managed to create 28 repositories with rich, varied content.

The Repository Service was developed through national and EU funding during 2012-2015 and aims to support knowledge producers (in the fields of culture, education, science and research) to organize, document and disseminate their content on an open access basis. Overarching goal of the project was to aggregate quality science and cultural content and disseminate through search portals and aggregators such as Europeana.

## KEYWORDS

Digital Repositories services; Data management and representation; DH, learning and pedagogy; Technologies for cultural heritage information

## INTRODUCTION

This proposal presents a robust cloud infrastructure of a Digital Repository Service (DRS) in Greece, developed by the National Documentation Centre (EKT) (<http://www.ekt.gr/en>) through the Operational Programme "Digital Convergence" (2011-2015). This Service

supports Digital Humanities (DH) as well as other content (e.g. science, research) through its lifecycle and is offered as SaaS (Software as a Service) public service to a wide number of research, cultural, educational and memory organisations along with e-learning and support services.

The main objective of the DRS is first to meet EU priorities for open access and economies of scale through the dissemination and re-use of large amounts of data and second to focus on the organisations needs regarding administration and dissemination of content with proper licensing.

The proposal presents the methodology used to plan and implement the whole service, its key characteristics and components, the functionality and interoperability with other systems and services, some indicative DH repository applications and the main conclusions obtained.

## METHODOLOGY

The methodology used to develop the DRS model built upon several factors: a) the National Documentation Centre long-term experience in developing data administration systems as well as collecting and maintaining high quality data for various purposes and users b) major past and current international practices and initiatives for managing and curating cultural and/or scientific data c) an extensive literature review on cultural ontologies and vocabularies used in similar initiatives and d) addressing community needs, the restrictions and the variety of the material.

The National Documentation Centre has been developing digital repository services and applications since the start of the century to support - mainly but not exclusively - the scientific community. “Helios” repository (<http://helios-eie.ekt.gr>) started off in 2007 comprising research and scientific content produced by researchers at the National Hellenic Research Foundation (<http://www.eie.gr>). DSpace is the platform of the repository but additional widgets and add-ons were developed by EKT to streamline the repository functionality. The schema is an application profile of Extended Dublin Core to incorporate scientific-related content. Pandektis (<http://pandektis.ekt.gr/>) is another digital repository of EKT, developed during 2000-2006, also running on DSpace. It uses Dublin Core and hosts rare Greek DH content.

Good international practices were also a major factor on planning the DRS. Similar implementations were studied[26][28][29] and major projects and standards were evaluated regarding the key implications when building cultural digital collections: the CARARE metadata schema[13] is quite characteristic on the cultural heritage sector since it drew on already developed standards such as: MIDAS, LIDO and CIDOC-CRM. These standards are a) highly conceptual b) event-centric[27] c) domain-specific[33] and d) as proven by the community analysis scarcely used (Fig. 2). Likewise there has been a discussion whether bibliographic conceptual models like the FRBR family[23][24][36][15] fit into all cultural heritage documentation settings or practically prove serviceable to users[15]. Major or smaller data ingestion initiatives like Europeana (<https://www.europeana.eu>)[21], DPLA (<https://dp.la>)[19] and the DM2E project (<https://dm2e.eu>)[1] were given particular importance in terms of their adopted schemes because of their holistic approach in managing heterogeneous collections.

To develop such a global service some key issues should be taken into account[31][32]: a) material heterogeneity, specifications and restrictions b) metadata semantic and structural viability and cohesion c) interoperability of data and systems d) data licensing e) data long-term preservation f) user/community studies and respective provisions.

To address the latter a carefully designed questionnaire was sent to various cultural organisations: libraries, museums, archives. The questionnaire would evaluate several aspects of the service: a) the community preparedness for it and most importantly for the sharing of

their material b) the variety of material owned by different organisations c) the possible restrictions and problems posed by related mentalities. The feedback provided us with valuable information on the community content needs and expectations.

Among 114 organisations, 90 had digitized material, 89 documented, while only 56 disseminated their content openly via internet. The most obvious observation of the results was diversity, primarily in terms of material (Fig. 1), as well as of documentation schemes and vocabularies used (Fig. 2). Moreover, most organisations stressed the difficulty to search and easily retrieve their material. Another key observation was their reluctance to open access and the reuse of their digital material.

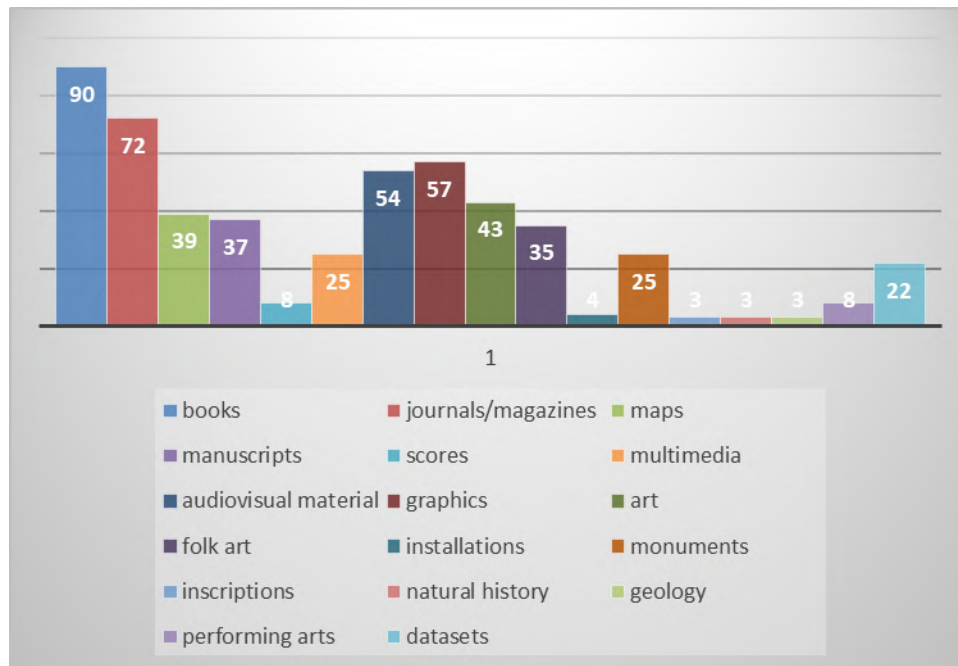


Figure 1. Material diversity among 114 organisations

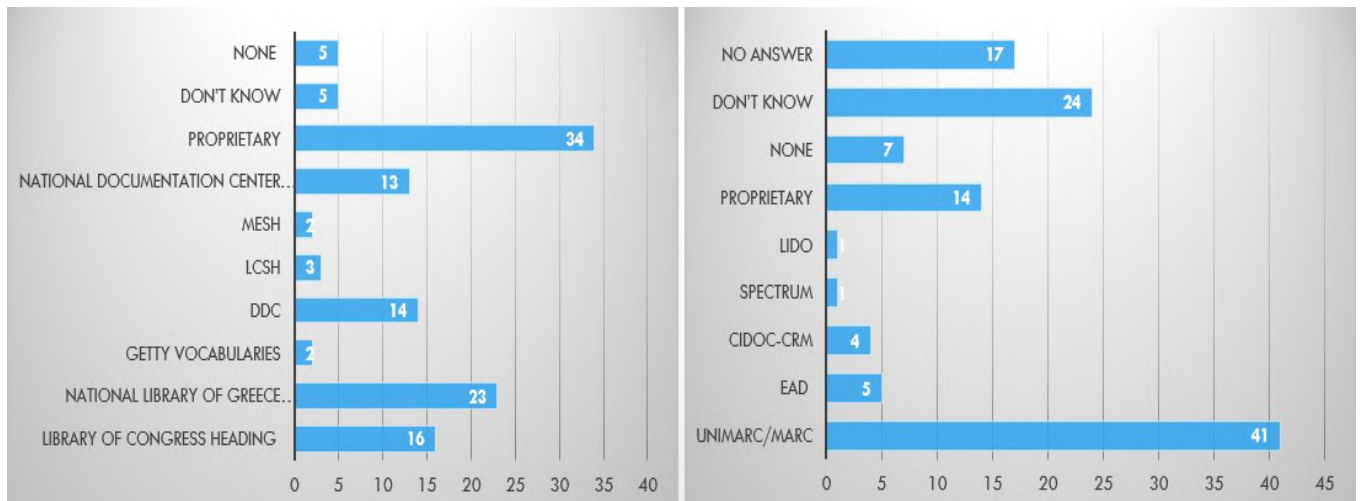


Figure 2. Ontologies and vocabularies used

## IMPLEMENTATION

Implementation of the DRS was based on the conclusions drawn over the methodology process. The service is structured by five major components (Fig. 3): a) the cloud infrastructure[37][18] which supports and boosts the whole service, providing also the necessary scalability b) the DRS schema that comprises the spinal column for metadata

structure<sup>3</sup> c) the Wizard Designing Tool (WDT) that collects and aligns all necessary specifications for creating a Digital Repository Application d) DSpace software for data management e) community engagement and support towards more customizable services.

The schema was based on the Qualified Dublin Core (DC), enriched with elements from different schemes. Apart from its popularity[30], flexibility[29] and interoperability with other schemes[27][12], DC is also OAI-PMH compatible which enlarges dissemination processes to other systems and services. For example EKT DRS schema has been mapped to EDM aggregator schema and data are disseminated to the Europeana portal. Its kinship to EDM would also allow for its extension into more classes in the future for describing concepts, places etc.

This approach was chosen because a global application profile should have a) sufficient rigidity for interoperability among data and b) adequate scalability and modularity to also include various types of material. In this sense the schema is a dynamic and modular corpus of elements that can adjust to different types of data needs and on the same time boosts up information retrieval by providing fixed elements with the same semantics to interoperate among different material forms<sup>4</sup> (e.g. type, creator, title, subject etc.).

Apart from DC, the schema was extended with properties of VRACore[35] and CDWA Lite[25] for the description of images, visual culture and artworks. For the description and evaluation of learning resources we used a set of elements from the Learning Object Metadata model (LOM)[22][28][30], while we consulted DataCite[16] to support datasets. Copyright issues were dealt with by providing information about the copyright holder, the copyright date and the CC license on the use of the digital object.

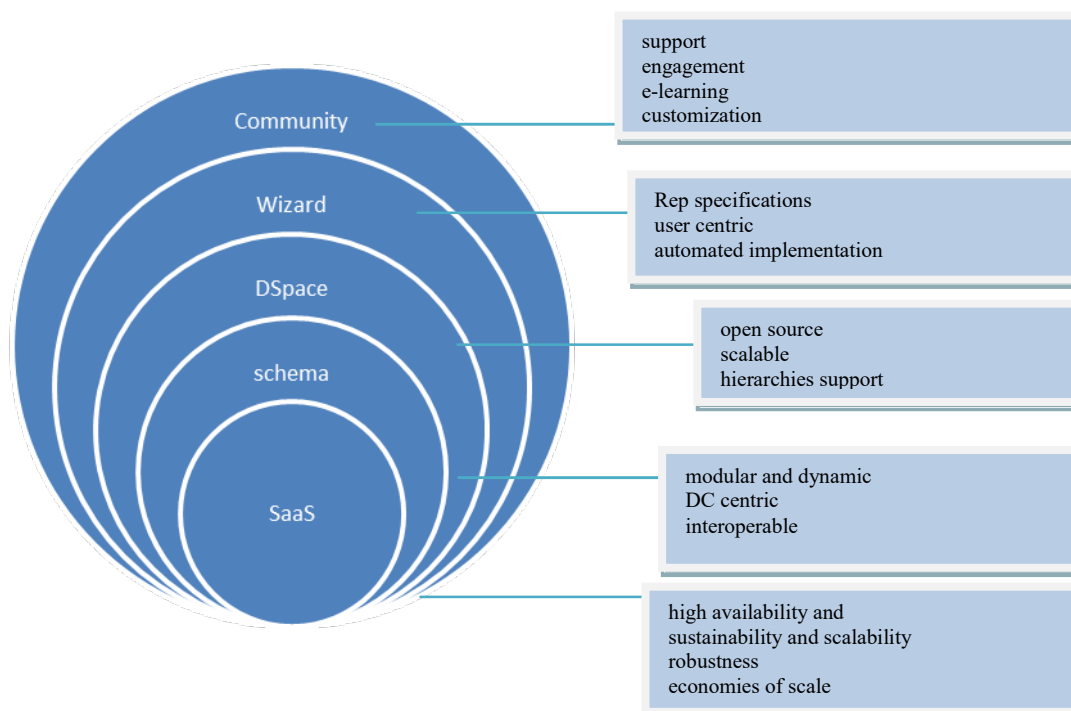


Figure 3. The Digital Repository Service model

<sup>3</sup> Issues of picking up appropriate established vocabularies for certain fields were also addressed (e.g Getty Art and Architecture Thesaurus).

<sup>4</sup> The DRS supports eight basic material types: text, graphical, maps, video, sound, scores, 3D objects and datasets.



DSpace software was selected as the data hosting and management system because it a) incorporates the well-established Dublin Core and its extended version[28] b) supports content dissemination processes with OAI-PMH protocol c) has flexibility in managing digital files as well as validation and migration files processes d) is open source and has a large interactive community e) supports content submission by various communities (researchers, scholars etc.) f) supports collection management (Communities--Collections—Items) as well as interconnection between resources supporting the complex relations of special material.

A very important module of the service is the Wizard Repository Designing Tool (WRDT) (Fig. 4). The tool is a functional dynamic procedure through which users can define all the specifications (communicational, metadata fields, vocabularies etc.) of their future Repository Application. As soon as someone fills in their specifications, the WRDT produces the Repository Application automatically. The tool provides also all the necessary instructions for the user to configure their future digital repository according to their needs.

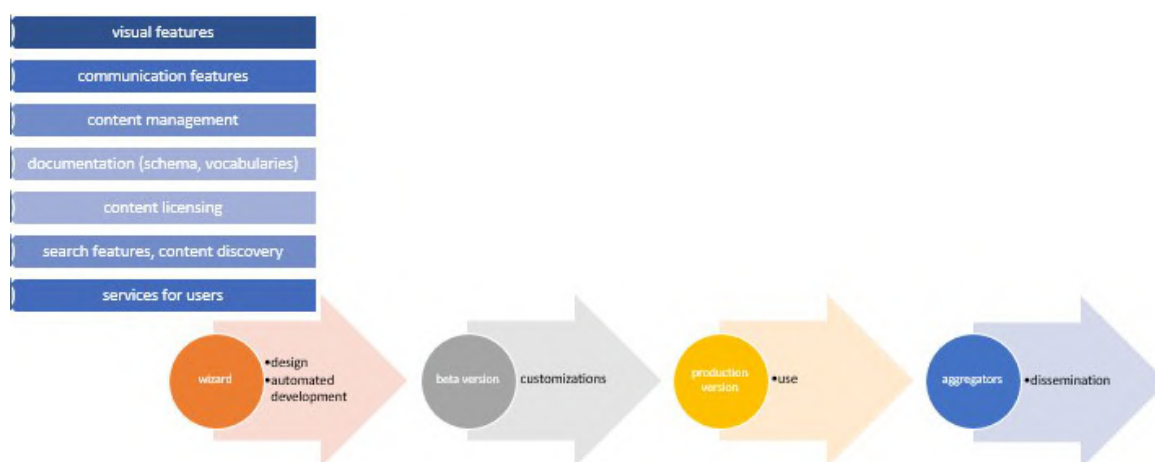


Figure 4. The Repository development process via the Wizard tool

Finally since the service is user-centric it could not live on without the community engagement and support. An extended e-Learning platform has been developed with appropriate material, guiding the user all the way through the service. More support material enriches the service feeding up its eKnowledgeBase and a stand-alone online service (e-HelpDesk) ensures that all problems are well cleared out.

## DRS APPLICATIONS

So far, DRS applications were directed to different kinds of organisations and were applied to various DH contexts and purposes: learning, research-academic etc.

In the pedagogy field DRS applications may be considered as key DH tools to instructors:

*Euterpe: Greek songs for schools - Digital Music Anthology:*  
<http://euterpe.mmb.org.gr/euterpe/?locale=en>

The musical repository incorporates a unique digital collection of songs in order to support music teaching in Greece. Songs are represented in music notation (scores, finale) and sound recordings and are accompanied by musicological information and teaching suggestions.

*Nicholas P. Goulandris Foundation - Museum of Cycladic Art: Learning Resources:* <http://repository-mca.ekt.gr/mca/>

This repository offers valuable cultural resources for learning purposes as well as learning resources, material and ideas for educational activities related to the Museum's permanent

collections and periodical exhibitions, which can be applied to school groups, families and children with disabilities.

In the cultural field the DRS model has developed some very interested applications from various cultural perspectives:

*Acropolis Restoration Service (YSMA)*: <http://repository-ysma.ekt.gr/ysma/about?locale=en>

The Repository disseminates original material from the restoration programs of the Acropolis of Athens. It is targeted to researchers, architects, archaeologists and scholars who are interested in the classical architecture and history, in Athenian Acropolis monuments or the management of cultural heritage. Data were migrated into the digital repository's infrastructure from an external database, after they were mapped to repository's metadata schema.

*The e-stories chestbox-Oral history from Serres*: <http://repository-e-stories.ekt.gr/>

The Repository records the experiential way in which a community of people recalls the past, setting the framework for reflection and mnemonic activity. The "e-stories chestbox" is a useful research tool for the multidisciplinary community of researchers involved in oral history, memory studies and social sciences.

*"TECHNI" Macedonian Art Society*: <http://repository-tehni-thes.ekt.gr/>

It showcases Thessaloniki's rich cultural production of six decades and offers an insight into the thinking and the work of major players in the arts and culture in the city.

## RESULTS

So far the DRS have produced 28 repositories applications. A key and quantifiable result is that through the DRS a significant volume of digital content has been deposited (more than 100.000 records and 5 million visits), which has later been aggregated in EKTs Aggregators, SearchCulture.gr and OpenArchives.gr. Part of it has been ingested by Europeana, as part of a national effort to increase good quality Greek content in the European Digital Library.

More qualitative results include knowledge transfer and sharing technical know-how with a community of users not specialized in knowledge and data management, but who were able to use the accompanying Wizard, eLearning and eKnowledgeBase service components to raise their documentation skills. In addition, the aggregated material offers possibilities for new projects and services. It has promoted collaborative actions for re-use of Digital Humanities content and applications for learning purposes.

The DRS schema is also currently being modified so as to incorporate EDM classes and is intended to connect to EKT vocabulary administration service Semantics.gr (<https://www.semantics.gr/>), which will boost its potential to maintaining national and international Authority entities and vocabularies.

## REFERENCES

- [1] Baierer, K., Dröge, E., Eckert, K., Goldfarb, D., Iwanowa, J., Morbidoni, C., & Ritze, D. 2017. *DM2E: A Linked Data source of Digitised Manuscripts for the Digital Humanities*. In: *Semantic Web*, 8(5), 733-745. DOI:10.3233/sw-160234. Retrieved January 2, 2019, from <http://www.semantic-web-journal.net/system/files/swj1299.pdf>
- [11] Bernadou, A., Champion, E., Dallas, C., Hughes, L., 2018. *Cultural Heritage Infrastructures in Digital Humanities*. Routledge
- [12] Carrasco, L., Silvana A. Borsetti Gregorio Vidotti. 2015. *Dublin core and CIDOC CRM harmonization*. In: *DCMI'15 Proceedings of the 2015 International Conference on Dublin Core and Metadata Applications*, São Paulo, Brazil — September 01 - 04, 2015, pages 198-200. Retrieved January 2, 2019, from <http://dcpapers.dublincore.org/pubs/article/view/3779>

- [13] CARARE Project. (2013). *CARARE Metadata Schema*. Retrieved January 2, 2019, from <http://pro.carare.eu/doku.php?id=support:metadata-schema>
- [14] CIDOC Conceptual Reference Model <http://www.cidoc-crm.org/>
- [15] Coyle, K. (2016). *FRBR: Before and after: A look at our bibliographic models*. Chicago: ALA Editions. Retrieved January 2, 2019, from <http://kcoyle.net/beforeAndAfter/978-0-8389-1364-2.pdf>
- [16] DataCite Metadata Working Group. 2017. *DataCite Schema*. Retrieved January 2, 2019, from [https://schema.datacite.org/meta/kernel-4.1/doc/DataCite-MetadadataKernel\\_v4.1.pdf](https://schema.datacite.org/meta/kernel-4.1/doc/DataCite-MetadadataKernel_v4.1.pdf)
- [17] DCMI Usage Board. (n.d.). DCMI: Home. Retrieved from <http://dublincore.org/>
- [18] Dehmlow, Mavodza, 2016. *Leading Change in the Age of the Cloud*, In: *Strategic Library*, Issue 15, pp 1-5. Retrieved 4 November, 2018, from <http://www.libraryspot.net/LPN/0416/Article%20%282%29.pdf>
- [19] DPLA MAP Working Group. 2017. *Introduction to the DPLA Metadata Application Profile*, version 5.0. Retrieved January 2, 2019, from <https://drive.google.com/file/d/1kMxXgFrGwu3i7LBLFOj6VZRQFuQzqkHk/view>
- [20] Duşa, A., Nelle, D., Stock, G., Wagner, G.G. 2014. *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences*. Bundesministerium für Bildung und Forschung. Retrieved 4 November, 2018, from <https://adriandusa.eu/publication/facingthefuture/>
- [21] Europeana. 2017. *EuropeanaDataModel-Mapping: Guidelines v2.4*. Retrieved January 2, 2019, from [https://pro.europeana.eu/files/Europeana\\_Professional/Share\\_your\\_data/Technical\\_requirements/EDM\\_Documentation/EDM\\_Mapping\\_Guidelines\\_v2.4\\_102017.pdf](https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Mapping_Guidelines_v2.4_102017.pdf)
- [22] IEEE STANDARDS ASSOCIATION. 2002, December 19. *IEEE 1484.12.1-2002 - IEEE Standard for Learning Object Metadata*. Retrieved January 2, 2019, from [https://standards.ieee.org/standard/1484\\_12\\_1-2002.html](https://standards.ieee.org/standard/1484_12_1-2002.html)
- [23] IFLA. 2017. *Functional Requirements: the FRBR Family of Models*. Retrieved January 2, 2019, from <https://www.ifla.org/ES/node/8914>
- [24] IFLA. 2017. *IFLA Library Reference Model (LRM)*. Retrieved January 2, 2019, from <https://www.ifla.org/publications/node/11412>
- [25] J. Paul Getty Trust ARTstor. 2006. *CDWA Lite: Specification for an XML Schema for Contributing Records via the OAI Harvesting Protocol*. Retrieved January 2, 2019, from [http://www.getty.edu/research/publications/electronic\\_publications/cdwa/cdwalite.pdf](http://www.getty.edu/research/publications/electronic_publications/cdwa/cdwalite.pdf)
- [26] JISC. 2009. *Kultur project document project sheet*. Retrieved January 2, 2019, from [http://kultur.eprints.org/Metadata\\_report\\_Final.pdf](http://kultur.eprints.org/Metadata_report_Final.pdf)
- [27] Kakali, C. et al. 2007. *Integrating Dublin Core metadata for cultural heritage collections using ontologies*. In: International Conference on Dublin Core and Metadata Applications, Singapore, August 27-31. Retrieved 4 November, 2019, from <http://dcpapers.dublincore.org/pubs/article/view/871>
- [28] Koutsomitropoulos, D., Alexopoulos, A., Solomou, G. Theodore S. Papatheodorou, T. 2010. *The Use of Metadata for Educational Resources in Digital Repositories: Practices and Perspectives*. In: D-Lib Magazine, volume 16, number 1/2. Retrieved January 2, 2019 from <http://www.dlib.org/dlib/january10/kout/01kout.html>
- [29] Morgan, M., Suhonos, MJ., Wang, F. 2013. *Digital Humanities and Metadata: Linking the Past to the Digital Future*. In: DCMI'13 Proceedings of the 2013 International Conference on Dublin Core and Metadata Applications, pages 206-208. Retrieved January 2, 2019 from <http://dcpapers.dublincore.org/pubs/article/view/3696>
- [30] Morgana, And., Baptista, A. (2015). *The use of Application Profiles and Metadata Schemas by Digital Repositories: Finding from a Survey*. In: DCMI International Conference on Dublin Core and Metadata Applications, São Paulo, Brazil, 1-4 September

2015. Retrieved January 2, 2019 from <http://dcpapers.dublincore.org/pubs/article/view/3770>
- [31] NISO Framework Working Group. 2007. *A Framework of Guidance for Building Good Digital Collections*. National Information Standards Organization (NISO).
- [32] Rahman, A. I. M. Jakaria & Francese, Enrico & Yilmaz, Muharrem & Beyene, Wondwossen. 2011. *Metadata practices in digital libraries*. In: International Seminar ‘Vision 2021: the role of libraries for building digital Bangladesh, At Dhaka, Bangladesh. Retrieved January 2, 2019, from [https://www.researchgate.net/publication/277597092\\_Metadata\\_practices\\_in\\_digital\\_libraries](https://www.researchgate.net/publication/277597092_Metadata_practices_in_digital_libraries)
- [33] Ronzino, Paola & Amico, Nicola & Niccolucci, Franco. 2011. Assessment and Comparison of Metadata Schemas for Architectural Heritage. In: CIPA Symposium. 4 November, 2018, from <http://eprints.rclis.org/11001/>, [https://www.researchgate.net/publication/259780242\\_Assessment\\_and\\_Comparison\\_of\\_Metadata\\_Schemas\\_for\\_Architectural\\_Heritage/download](https://www.researchgate.net/publication/259780242_Assessment_and_Comparison_of_Metadata_Schemas_for_Architectural_Heritage/download)
- [34] Simons, N., Richardson, J. 2013. *New Content in Digital Repositories: The Changing Research Landscape*. Woodhead Publishing Limited.
- [35] VRACORE: a data standard for description of images and works of art and culture. <https://www.loc.gov/standards/vracore/>
- [36] Žumer, M. and Riva, P. 2017. *IFLA LRM - finally here*. In *Proceedings of the 2017 International Conference on Dublin Core and Metadata Applications (DCMI'17)*. Dublin Core Metadata Initiative 13-23. Retrieved January 2, 2019, from <http://dcpapers.dublincore.org/pubs/article/view/3852/2037>
- [37] Μπάρτζη, Κ., Αλεξία Παναγοπούλου, Α., Λαγούδη, Ε., Βασιλογαμβράκης, κ.ά. 2017. *Τεχνολογίες αιχμής και υπηρεσίες νέφους: ένα βιώσιμο μοντέλο υπηρεσιών για βιβλιοθήκες & φορείς περιεχομένου*. In: 23ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών. Retrieved 4 November, 2018, from <https://helios-eie.ekt.gr/EIE/handle/10442/15618>

# Didattica universitaria e ambienti digitali

Chiara Pancioli, Anita Macauda

University of Bologna, Italy, chiara.pancioli@unibo.it, anita.macauda@unibo.it

## ABSTRACT

Secondo una *prima* istanza, la didattica può essere intesa innovativa quando sa creare una dialettica tra complessità del contesto e soluzioni metodologiche efficaci, sia in termini di mediazione, sia di sostenibilità dei processi. In *seconda* istanza, la didattica innovativa si lega al concetto di ambiente di apprendimento, in quanto il rinnovamento delle pratiche didattiche richiede un approccio sistemico, attento non solo alle metodologie impiegate, ma anche a tutti gli elementi che configurano l'evento formativo. In questa prospettiva il concetto di ambiente richiama quello di ecosistema fisico-digitale entro il quale si viene a collocare l'esperienza educativa (Pancioli 2017). Infatti, oggi il confine tra ambienti fisici e digitali tende sempre più ad assottigliarsi e in tal senso occorre un ripensamento dell'interazione uomo-informazione. In *terza* istanza, la didattica innovativa, in una prospettiva multimodale dell'esperienza educativa, fa riferimento all'uso integrato di differenti linguaggi (visivo, grafico, sonoro, video, corporeo), con un'attenzione particolare ai processi di acquisizione e rielaborazione delle conoscenze.

Le diverse istanze finora delineate contribuiscono alla definizione di un *modello* didattico necessario per progettare, realizzare e valutare i processi di insegnamento/apprendimento in uno specifico ambiente, in vista del conseguimento di determinati obiettivi. La dimensione tecnologica si deve sviluppare parallelamente e sinergicamente alle ipotesi di lavoro che sono alla base della formulazione di progetti formativi (Baldacci, 2004). Partendo da questo quadro generale, l'articolo prende in esame alcune strategie d'insegnamento sperimentate, in rapporto ad ambienti sia fisici sia digitali, in alcuni corsi di studio dell'Università di Bologna.

## PAROLE CHIAVE

Didattica innovativa, ecosistema fisico-digitale, flipped classroom, cooperative learning, mediazione didattica.

## INTRODUZIONE

Il contributo vuole riflettere su alcuni aspetti significativi che permettono di descrivere gli elementi necessari per innovare la didattica universitaria. Un interesse particolare sembra oggi indirizzarsi verso un campo specifico, riguardante l'innovazione dei setting d'apprendimento: se da un lato, esso offre l'opportunità di ripensare alla didattica come ambito di azioni complesso, dall'altro rivela aspetti di incertezza rispetto all'identificazione degli elementi necessari per attuare il cambiamento. Dopo aver individuato le principali istanze secondo cui la didattica si definisce innovativa, questo contributo prende in esame alcune strategie d'insegnamento (*flipped classroom, cooperative learning, project based learning,...*), in rapporto sia ad ambienti fisici sia ad ambienti digitali, sperimentate in modalità *blended learning* in alcuni corsi di studio dell'Università di Bologna.

## STATO DELL'ARTE

La didattica innovativa può essere definita come ambito di riflessione per individuare specifiche categorie empiriche degli eventi educativi. Si esplica all'interno di differenti contesti formativo-culturali, caratterizzati da logiche ecosistemiche in cui far dialogare gli

ambienti formali e informali dell'apprendimento. La didattica innovativa porta quindi a ripensare alle azioni e alle soluzioni più significative richiamando strategie operative che trasformino l'allestimento dei setting formativi anche alla luce degli sviluppi tecnologici. La didattica così intesa si struttura attraverso fasi di natura progettuale, attuativa, valutativa, di negoziazione per sostenere processi di acquisizione della conoscenza, in relazione anche al bagaglio motivazionale, esperienziale e valoriale della/e persona/e che apprendono.

In *prima* istanza la didattica si definisce innovativa quando sa creare una dialettica tra complessità del contesto e soluzioni metodologiche efficaci sia in termini di mediazione, sia di sostenibilità dei processi. Nello specifico della mediazione, essa si realizza mediante processi di metaforizzazione (Rézeau 2002; Damiano 2013; Rossi 2016) che permettono di operare semplificazioni sostituendo il reale con modelli e/o simboli. Rispetto invece agli elementi di sostenibilità, intendendo per sostenibilità la possibilità di misurare il successo di un processo caratterizzato da innovazione, gli studi specifici più recenti (Trentin, 2008; Panciroli 2008; Eickelmann, 2011; Ferrari, 2017a; 2017b) collocano il successo delle pratiche didattiche in uno spazio metariflessivo, in cui monitorare diverse dimensioni strettamente e mutuamente interrelate (socio-culturale, economica, professionale, contenutistica, formale-informale).

In *seconda* istanza, connesso al significato di didattica innovativa, vi è il concetto di ambiente di apprendimento, a partire dalla convinzione che il rinnovamento delle pratiche richieda un approccio sistemico, attento non solo alle metodologie impiegate dall'insegnante o alla gestione della relazione con gli allievi, bensì a tutti gli elementi che configurano l'evento formativo. In questa prospettiva il concetto di "ambiente" richiama il concetto di ecosistema fisico-digitale entro il quale avviene l'esperienza formativa (Rosati 2013; Mura, 2014) e in cui riconoscere gli allievi come attori protagonisti, incoraggiare il loro impegno attivo e promuovere una auto-comprensione della propria esperienza, basarsi sulla natura sociale dell'apprendimento e incoraggiare attivamente il lavoro di gruppo.

In *terza* istanza, la didattica innovativa, in una prospettiva multimodale dell'esperienza educativa, fa riferimento all'uso integrato di differenti linguaggi (visivo, grafico, sonoro, video, corporeo...) nei processi di acquisizione e rielaborazione delle conoscenze. Nello specifico del visivo, l'impiego delle immagini richiama la possibilità di stimolare e attivare prenoscenze adeguate agli oggetti di apprendimento, di sviluppare la capacità di problematizzare, di mettere in connessione gli elementi della propria conoscenza in contesti complessi (Calvani 2011; Cardarello, Contini, 2012; Landriscina, 2012; Cicalò 2016; Corazza 2017; Panciroli, Luigini, 2018; Macaudo 2018).

Le tre istanze finora delineate contribuiscono alla definizione di un modello didattico necessario per progettare, realizzare e valutare i processi di insegnamento/apprendimento in uno specifico ambiente, in vista del conseguimento di determinati obiettivi. In tal senso, la dimensione tecnologica si deve sviluppare parallelamente e sinergicamente alle ipotesi di lavoro che sono alla base della formulazione di progetti formativi (Baldacci, 2004; Laurillard 2014).

## **OBIETTIVI DELLA RICERCA**

Il principale obiettivo della ricerca è stato quello di individuare, secondo un approccio quantitativo e qualitativo, gli elementi significativi di una didattica innovativa in un'ottica ecosistemica in cui si intrecciano in modo sempre più naturale spazi didattici fisico-virtuali e media differenti e in cui il contesto gioca un ruolo decisivo per il raggiungimento di specifici obiettivi (Rosati 2013; Rivoltella 2017). La possibilità infatti di utilizzare molteplici linguaggi e di metterli in rete, grazie alla digitalizzazione, può portare alla creazione di traiettorie plurime. La scrittura in rete, così come le riletture multiple garantiscono un continuo confronto tra l'agito e l'agire, tra la propria prospettiva e la prospettiva altrà (Rossi 2010).

Come evidenzia Rossi, molte delle tecnologie del web garantiscono a più soggetti di collaborare e di costruire artefatti composti da frammenti, ognuno dei quali mantiene la propria autonomia e identità ma che, essendo inseriti in un sistema condiviso, possono dialogare.

La ricerca è stata condotta attraverso la sperimentazione di alcune specifiche strategie, in rapporto ad ambienti sia fisici sia digitali, nell'ambito dei corsi di "Iconografia e iconologia" (laurea magistrale a ciclo unico in Scienze della Formazione Primaria), di "Teorie e strumenti della mediazione didattica" (laurea magistrale in Pedagogia) e di "Didattica museale" (laurea triennale in Educatore sociale e culturale) del Dipartimento di Scienze dell'Educazione dell'Università di Bologna.

## **SPERIMENTAZIONE**

La sperimentazione ha coinvolto complessivamente 486 studenti frequentanti nell'anno accademico 2017/18. Nello specifico: 252 studenti del corso di "Iconografia e iconologia"; 70 studenti del laboratorio di "Iconografia e iconologia"; 94 studenti del corso di "Teorie e strumenti della mediazione didattica"; 58 del corso di "Didattica museale".

### **Le strategie sperimentate**

I corsi d'insegnamento sono stati progettati tenendo in riferimento alcune specifiche architetture didattiche: informativo-ricettiva, metacognitivo-creativa e socio-relazionale.

Nell'ambito di un'architettura *informativo-ricettiva* basata appunto sulla trasmissione delle informazioni essenziali (conoscenze e lessico di base, parole chiave,...), è stato adottato un tipo di lezione multimodale supportata da immagini (quadri, fotografie, disegni, illustrazioni,...), video (documentari, spezzoni cinematografici,...), testi (approfondimenti, ricerche,...) che sono serviti da stimolo per avviare il confronto su determinati concetti teorici.

Rispetto all'architettura *metacognitivo-creativa*, sono stati forniti agli studenti risorse e supporti che hanno permesso una parziale pre-strutturazione dell'informazione e una costruzione attiva della conoscenza. Nello specifico, attraverso la lezione euristica o partecipata che ha un carattere prettamente dialogico, il docente ha alternato brevi esposizioni alla formulazione di domande, in base a un continuo scambio-confronto di idee con gli studenti. Questi non hanno avuto solo il ruolo di ascoltatori, ma hanno preso parte all'elaborazione/costruzione dei contenuti in riferimento alle continue problematizzazioni sollevate dal docente (*problem solving*) secondo un approccio metacognitivo. In questo tipo di architettura rientra anche il *role playing*, una strategia didattica basata sulla simulazione di atteggiamenti e comportamenti adottati in specifici contesti lavorativi; questa attività ha portato gli studenti a percepire specifici aspetti che emergono nei rapporti interpersonali e che risultano difficili da formalizzare verbalmente.

L'architettura *socio-relazionale* ha presupposto l'acquisizione da parte degli studenti di un alto grado di autonomia nella formulazione di un proprio pensiero, da condividere con gli altri nell'ambito della realizzazione di un progetto comune. Nello specifico si è fatto riferimento a un tipo di apprendimento basato sull'interazione tra pari (*peer learning, peer tutoring*) e sull'apprendimento di gruppo, in cui è stato dato ampio margine all'apprendimento cooperativo (*cooperative learning*).

Queste tre architetture didattiche sono state sviluppate in modalità *blended learning* e nello specifico sono state supportate da una piattaforma moodle dedicata a ciascun insegnamento, secondo una strutturazione delle lezioni distanza-presenza-distanza (tabella 1). Questa modalità ha orientato i corsi verso un formato didattico proprio della *flipped classroom* (Bergmann, Sams 2012; Maglioni, Biscaro, 2014; Slomanson 2014; Bevilacqua 2018),

letteralmente “classe capovolta”. Infatti, le sequenze di lavoro sono state ribaltate: l’attività di studio a casa ha preceduto l’attività di approfondimento e riflessione critica svolta in aula. Lo studio a casa è stato affidato sia a contenuti didattici digitali (testi, video e altri materiali multimediali) organizzati in forma di Learning Object (LO), sia ad attività cooperative (*e-tivity*) erogate su piattaforma Moodle. L’insegnamento capovolto ha trasformato così l’aula in uno spazio di confronto, approfondimento, applicazione e valutazione, attraverso la realizzazione di attività collaborative (*cooperative learning, peer tutoring*), esperienze (*role playing*), riflessioni e discussioni (*problem solving*).

La sperimentazione descritta ha evidenziato un cambio di paradigma nell’utilizzo delle tecnologie: progettate, a partire dagli anni Sessanta, come mezzi, vengono reinterpretate, negli anni Ottanta-Novanta, come ambienti. La prospettiva ambientale delle tecnologie suggerisce appunto una maggiore pervasività, proponendo l’idea di un mezzo-ambiente in cui l’uomo è inserito e in cui può fare delle cose all’interno di uno specifico contesto d’azione (Rivoltella, 2014). Il confine tra spazio fisico dell’aula e spazio digitale della piattaforma tende pertanto ad assottigliarsi, fin quasi a scomparire. Questo determina un intreccio sempre più inestricabile fra ambienti eterogenei, dando luogo ad un vero e proprio ecosistema e – con particolare riferimento all’intreccio fisico-digitale – a un *ubiquitous ecology*. Le ecologie ubique sono sistemi emergenti in cui vecchi e nuovi media, ambienti fisici e digitali sono progettati, realizzati e fruiti come un tutt’uno omogeneo con la stessa pervasività dello *ubiquitous computing* (Kuniawsky 2010; Resmini, Rosati 2011).

Fase	ATTIVITÀ	STRUMENTI	ABILITÀ/COMPETENZE	ESITI
I	Lezioni frontali, dibattito (presenza)	Testi, presentazioni multimediali	Ascoltare, dialogare	Discussione in classe
II	Presentazione di immagini, video, casi studio sui contesti educativi (presenza)	Museo virtuale (MOdE), siti web	Classificare i modelli di didattica, riconoscere i differenti profili professionali	Individuazione differenti modelli di didattica e ruolo docente/educatore all’interno di differenti contesti
III	Studio dei contenuti di approfondimento (distanza)	Piattaforma Moodle	Approfondire	Contestualizzate le informazioni/conoscenze attraverso l’analisi, la memorizzazione e la comparazione
IV	Visione di un documentario d’osservazione e riflessione (distanza)	Canale YouTube	Osservare in modo professionale: riconoscere e decodificare	Riflessione preparatoria alla fase del brainstorming
V	Brainstorming (presenza)	Caso studio	Riflettere in modo critico: formulare ipotesi	Mappa concettuale
VI	Lavoro collaborativo a piccoli gruppi (presenza)	Scheda di progettazione, tabella per riflessione guidata	Sintetizzare e confrontare Capacità di problem solving	Tabelle compilate
VII	Compilazione questionario (distanza)	Questionario (piattaforma moodle)	Meta-riflettere	Formalizzazione delle conoscenze
VIII	Restituzione esiti lavoro di gruppo e questionario (presenza)	Report	Comunicare Presentare oralmente	Dibattito e formulazione di nuove strategie

Tabella 1. Sintesi delle principali attività sperimentate in presenza e a distanza



## Gli strumenti della ricerca

Alla metà e alla fine del corso sono stati somministrati agli studenti due questionari valutativi in forma anonima e informatizzata. Il *primo* questionario, somministrato *in itinere* e articolato su quattro domande a risposta aperta, era orientato a comprendere il livello di significatività delle strategie didattiche sperimentate:

1. Sei finora soddisfatto del corso che stai frequentando?
2. Le metodologie presentate a lezione sono secondo te adeguate al profilo professionale del corso di studi da te scelto?
3. La metodologia blended learning è funzionale all'apprendimento dei contenuti?
4. Secondo il tuo parere ci sono dei punti di forza e delle criticità del corso che stai frequentando?

Il *secondo* questionario di valutazione sommativa, distribuito nell'ultima settimana di ciascun corso, era costituito da domande a risposta chiusa e aperta relative all'organizzazione complessiva (tempi, spazi integrati, strumenti), al gradimento rispetto alle metodologie adottate e alle conoscenze acquisite (chiarezza dei temi trattati, acquisizione di nuovi contenuti, trasferibilità, interesse generale). Di questo questionario, sono stati analizzati nello specifico quattro quesiti utili per capire i principali aspetti riguardanti la didattica innovativa:

1. Il docente stimola / motiva l'interesse verso la disciplina?
2. Sei complessivamente soddisfatto/a di come è stato svolto questo insegnamento?
3. Segnala gli aspetti positivi del corso seguito.
4. Segnala eventuali difficoltà o problemi riscontrati durante il corso seguito.

## RISULTATI OTTENUTI

Le risposte al *primo* questionario fornite dagli studenti hanno evidenziato un elevato grado di soddisfazione. In generale, è emerso come una didattica innovativa incentrata su differenti strategie possa facilitare lo sviluppo del pensiero critico (92% delle risposte degli studenti), l'interesse all'ambito di studio (90%) e la motivazione alla partecipazione e frequenza ai corsi d'insegnamento (88%). In particolare, l'analisi dei dati raccolti ha evidenziato: un forte coinvolgimento emotivo (stupore, entusiasmo, eccitazione); un significativo coinvolgimento metacognitivo (riflessioni critiche e problematizzanti rispetto ai dati osservati); l'auto-riflessione sulle proprie competenze. In particolare, le risposte alla domanda n. 3 (*La metodologia blended learning è funzionale all'apprendimento dei contenuti?*) hanno evidenziato molti aspetti positivi riguardanti il processo di apprendimento attraverso la piattaforma Moodle: la possibilità di uno studio più approfondito, di una maggiore riflessione e personalizzazione della conoscenza, oltre alla possibilità di poter svolgere attività che non possono essere realizzate in aula. Il limite riscontrato (7%) riguarda il fatto che questi ambienti richiedono un grado di autonomia elevata, una motivazione allo studio costante e puntuale per poter raggiungere una preparazione adeguata allo svolgimento delle attività che si sviluppano in presenza. A lezione, infatti, il docente sollecita gli studenti rispetto a determinati contenuti approfonditi individualmente in piattaforma.

Nel *secondo* questionario sono emersi pochi elementi di criticità (4% delle risposte degli studenti) riguardanti il disorientamento per un approccio ai contenuti aperto a molteplici stimoli che ha presupposto un'attività di riflessione e un tempo di studio maggiori rispetto a quelli richiesti dalla tradizionale lettura del manuale. Rispetto invece ai punti di forza (fig. 1),

le risposte hanno segnalato come aspetti positivi dei corsi: la metodologia didattica (25%); le strategie (25%); l'orientamento al profilo professionale (20%); gli spazi (12%); i linguaggi e i materiali (12%); la relazione tra docente e studenti (6%).

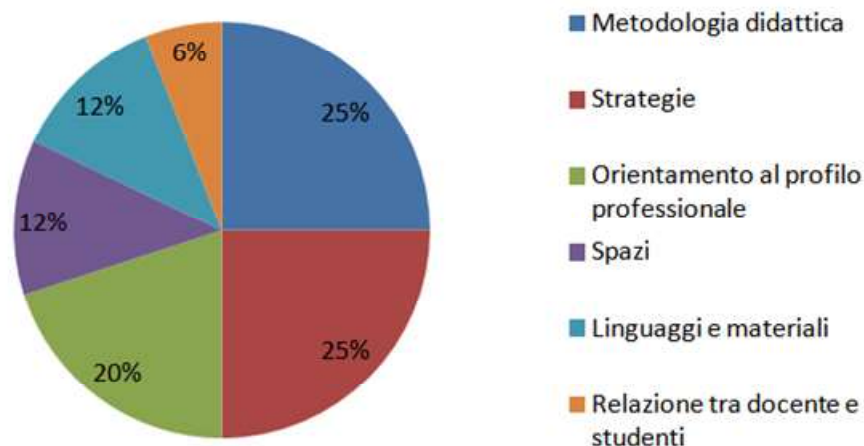


Fig. 1. Aspetti positivi dei corsi

La metodologia didattica su cui si è basata la sperimentazione è stata definita sul concetto di co-costruzione della conoscenza (conoscenza multiforme, connessioni logiche, problematizzazioni...).

I principali elementi di innovatività sono così sintetizzabili:

- - l'utilizzo di linguaggi e strategie differenti;
- - l'individuazione di obiettivi a diversi livelli (di base, intermedi, superiori convergenti, superiori divergenti);
- - il ruolo attivo degli studenti nella costruzione della conoscenza;
- - i differenti ruoli del docente: trasmissivo, di scaffolding, di sollecitatore, di tutoring, di connessione logica dei saperi;
- - l'utilizzo di strumenti e risorse specifiche (risorse multimediali, tabelle, mappe concettuali, piattaforma moodle; laboratori di didattica).

In particolare, la sperimentazione ha perseguito i seguenti obiettivi: *monocognitivi* (memorizzare, riconoscere, descrivere, esemplificare le principali strategie didattiche, tradizionali e di apprendimento attivo); *metacognitivi* (analizzare, confrontare e sintetizzare); *superiori-trasversali* (riconoscere il problema, tentare soluzioni rispetto a un setting didattico, inventare per analogia e formulare situazioni nuove); *socio-relazionali* (esplicitare problemi aperti; ascoltare delle esperienze personali, valorizzare gli stili d'apprendimento).

## BIBLIOGRAFIA

- [38] Baldacci M. 2004. *I modelli dell'insegnamento nell'epoca della società conoscitiva*. In Id. (Ed.), *I modelli della didattica* (pp. 13-59). Carocci. Roma. ISBN: 9788843029792
- [39] Bergmann J., & Sams A. 2012. *Flip your classroom: reach every student in every class every day*. International Society for Technology in Education.
- [40] Bevilacqua A. 2018. *Flipped learning in ambito universitario. Presupposti e indicazioni pedagogico-didattici tra implementazione e ricerca*. Pensa Multimedia. Lecce. ISBN: 978-88-6760-592-7

- [41] Calvani A. (ed.). 2007. *Fondamenti di didattica: teoria e prassi dei dispositivi formativi*. Carocci. Roma. ISBN: 978-8843042487
- [42] Calvani A. (ed.). 2011. *Principi di comunicazione visiva e multimediale. Fare didattica con le immagini*, Carocci, Roma. ISBN: 9788843061709
- [43] Cardarello R., Contini A. (ed.). 2012. *Parole, immagini metafore. Per una didattica della comprensione*, Edizioni Junior, Azzano San Paolo. ISBN: 9788884346278
- [44] Cicalò E. 2016. *Intelligenza grafica*, Aracne. Roma. ISBN:8854887234
- [45] Clark RC. 2000. *Four architectures of instruction*. *Perform Improvement*. 39(10): pp. 31-38. DOI: 10.1002/pfi.4140391011.
- [46] Corazza L. 2017. *Apprendere con i video digitali. Per una formazione online aperta a tutti*. FrancoAngeli. Milano. ISBN: 9788891762719
- [47] Damiano E. 2013. *La mediazione didattica. Per una teoria dell'insegnamento*. Franco Angeli. Milano. ISBN: 9788820450595
- [48] Eickelmann B. 2011. Supportive and hindering factors to a sustainable implementation of TIC in schools. *Journal for Educational Research Online*, 3, 75-103.
- [49] Ferrari L. 2017a. *Implementazione e sostenibilità di nuove tecnologie in El Salvador*. L'integrazione scolastica e sociale. 2, 158-165.
- [50] Ferrari L. 2017b. *Il digitale a scuola. Per una implementazione sostenibile*. FrancoAngeli. Milano. ISBN: 9788891753274
- [51] Landriscina F. 2012. *Didattica delle immagini: dall'informazione ai modelli mentali*. *Form@re*. Open Journal per la formazione in rete, vol. 12, 80: 27-34. DOI: 10.13128/formare-12610
- [52] Laurillard D. 2014. *Insegnamento come scienza della progettazione: costruire modelli pedagogici per apprendere con le tecnologie*, FrancoAngeli. Milano. ISBN: 9788820462796
- [53] Kuniawsky M. 2010. *Smart Things: Ubiquitous Computing User Experience Design*. Morgan Kaufmann. Burlington, Massachusetts. ISBN-13: 978-0123748997
- [54] Macaudo A. 2018. *Arte e innovazione tecnologica per una didattica immersiva*. FrancoAngeli. Milano. ISBN: 9788891768940
- [55] Maglioni M., Biscaro F. 2014. *La classe capovolta: innovare la didattica con la flipped classroom*. Erickson. Trento. ISBN: 9788859004882.
- [56] Merrill M.D. 2002. *First principles of instruction. Educational technology research and development*. 50(3), 43–59.  
<http://csapoer.pbworks.com/f/First+Principles+of+Instruction+>
- [57] Nigris E., Negri S.C., Zuccoli F. 2007. *Esperienza e didattica. Le metodologie attive*. Carocci. Roma. ISBN: 9788843042517
- [58] Panciroli C. 2008. *E-learning e learning-e. Riflessioni sulla formazione*. *Ricerche di Pedagogia e Didattica*, vol. 3. ISSN 1970-2221
- [59] Panciroli C. 2016. *Le professionalità educative tra scuola e musei*. Guerini. Milano. ISBN: 9788881074006
- [60] Panciroli C. 2017. *Ecosistemi digitali*. In L. Corazza (ed.). *Apprendere con i video digitali. Per una formazione online aperta a tutti* (pp. 13-32). FrancoAngeli. Milano. ISBN: 978-8891762719
- [61] Panciroli C. 2018. *Didattica scolastica e ambienti digitali per l'apprendimento*. In S. Ulivieri (ed.). *Le emergenze educative della società contemporanea. Progetti e proposte per il cambiamento*. Pensa Multimedia. Lecce. ISBN: 978-88-6760-584-2
- [62] Panciroli C., Corazza L., Vignola P., Marcato E., Leone D. 2018. *Didattica innovativa. Soluzioni efficaci per contesti complessi*. *Form@re-Open Journal per la formazione in rete*, 18, 2, pp.116-129. ISSN: 1825-7321. DOI: <http://dx.doi.org/10.13128/formare-23349>

- [63] Panciroli C., Luigini A. 2018. *Ambienti digitali per educare all'arte e al patrimonio*. FrancoAngeli. Milano. ISBN: 9788891773333
- [64] Reigeluth C.M. 1999. *The Elaboration Theory: Guidance for Scope and Sequence Decisions*. In C.M., Reigeluth (ed.), *Instructional Design theories and models: a new paradigm of Instructional Design* (pp. 425-453). Mahwah, NJ: Lawrence Erlbaum Associates.
- [65] Resmini A., Rosati L. 2011. *Pervasive Information Architecture: Designing Cross-Channel User Experiences*. Morgan Kaufmann. Burlington, Massachusetts. ISBN: 978-0123820945
- [66] Rézeau J. 2002. *Médiation, médiatisation et instruments d'enseignement : du triangle au «carré pédagogique»*, ASp la revue du GERAS, 35-36. URL : <http://journals.openedition.org/asp/1656>; DOI : 10.4000/asp.1656
- [67] Rivoltella P.C. 2014. *La previsione. Neuroscienze, apprendimento, didattica*. Brescia. Editrice La Scuola. ISBN: 978-8835038986
- [68] Rivoltella P. C. 2017. *Media Education. Idea, metodo, ricerca*. ELS La Scuola. Brescia. ISBN: 9788826500232
- [69] Rosati L. 2013. Ecosistemi fisico-digitali: progettare e produrre contenuti per i media integrati. *Problemi dell'informazione*, 1, 12-24. DOI: 10.1445/74592
- [70] Rossi P.G. 2010. *Tecnologia e costruzione di mondi: post-costruttivismo, linguaggi e ambienti di apprendimento*. Armando, Roma. ISBN: 8860816386
- [71] Rossi P.G. 2016. *Gli artefatti digitali e i processi di mediazione didattica*. *Pedagogia Oggi*, 2, 11-26.
- [72] Slomanson W. R. 2014. *Blended Learning: A Flipped Classroom Experiment*. *Journal of Legal Education*, vol. 64, No. 1, 93-102. <http://www.jstor.org/stable/24716075>.
- [73] Trentin G. 2008. *La sostenibilità didattico-formativa dell'e-learning: social networking e apprendimento attivo*. FrancoAngeli. Milano. ISBN: 9788846491343

# The Use of Blockchain for Digital Archives: Challenges and Perspectives

Clara Bacciu, Angelica Lo Duca Andrea Marchetti

IIT - CNR, Italy, [clara.bacciu@iit.cnr.it](mailto:clara.bacciu@iit.cnr.it), [angela.loduca@iit.cnr.it](mailto:angela.loduca@iit.cnr.it), [andrea.marchetti@iit.cnr.it](mailto:andrea.marchetti@iit.cnr.it)

## ABSTRACT

Over the last few years, the exploitation of blockchain technologies is increasing across different sectors, such as health and economy. A field that still remains little explored is that of digital archives owned by public administrations and other authoritative organizations. This paper investigates challenges and perspectives about the use of blockchain technologies for digital archives.

We also present an application scenario, which employs blockchain technology for the registration of minor artworks. The described platform deals with problems of artwork counterfeiting, loss and subsequent discovery, natural disasters and traceability.

## KEYWORDS

Blockchain; Digital Archives; Digital Objects.

## INTRODUCTION

Recently, the diffusion of applications based on blockchain technology [8, 10] has been increasing rapidly. The original focus of these technologies concerned cryptocurrencies (i.e., Bitcoin), but is shifting to finance and business in general, and is being extended progressively for a variety of applications in healthcare, government, Internet of Things, entity and assets management. A field which still seems to remain little explored is Cultural Heritage. We would like to investigate on the possible challenges and benefits derived by the application of blockchain technologies to this field, and, more in general, to digital archives.

Literature about this topic is still little. The ARCHANGEL project<sup>5</sup> will implement trusted archives of digital public records on a distributed ledger. The project has started last year and it will end by the end of 2018. Memory Matrix<sup>6</sup> is an ongoing project aiming at inserting information about Cultural Heritage in the Bitcoin network. The Kapu project<sup>7</sup> is another ongoing activity, which tries to create a blockchain for archeology.

The potential benefits introduced by the use of blockchain for digital archives are essentially the following:

a) replication, which prevents loss of information; b) documented provenance; c) immutability and resistance to unauthorized changes.

## AN OVERVIEW OF BLOCKCHAIN

A blockchain is a particular implementation of a Distributed Ledger (DL). A DL is essentially a database, which is shared among different nodes of a network. In practice, all the nodes of the network share the same copy of the database and any change made on a node, is replicated to all the other nodes in few minutes and, in some cases, even in few seconds. A DL can be

---

<sup>5</sup> <http://gow.epsrc.ac.uk/NGBOViewGrant.aspx?GrantRef=EP/P03151X/1>

<sup>6</sup> <http://www.memorymatrix.org>

<sup>7</sup> <https://www.kapu.one>

public (as opposite of private) if any node can read the content, and permissionless (as opposed of permissioned) if any node can write content.

The protocol for the first functioning blockchain was introduced in 2008 to support the digital cash Bitcoin [6], and implements the ledger as a chain of blocks. Each block contains data, a timestamp and a cryptographic hash of the previous block. This way the integrity of the information stored in the blockchain is protected through a security system based on cryptography.

With respect to a standard database, a blockchain is an append-only register. This means that information can only be added to the database, but it cannot be removed. Modifications to the stored data can be done by re-uploading a new version of the data.

A distributed consensus algorithm is used to decide which updates to the ledger are to be considered valid. New participants (nodes) can start collaborating to the maintenance of the repository by following this algorithm. There is no need of a central authority or trust between nodes; the consensus algorithm and cryptography grant the correctness of data even in presence of some malicious nodes.

Among the most important blockchain protocols are: the above mentioned Bitcoin [6], Ethereum [9], Hyperledger [3], Ripple Transaction Protocol (RTXP) [7].

## THE CONCEPT OF DIGITAL ARCHIVE

Archives contain records or collections of records that need long-term or even permanent preservation for their cultural, historical, or evidentiary value. In digital archives, a record can be anything holding a piece of information in the form of digital object.

The creation, management and use of a digital archive is not an easy task. ARMA International<sup>8</sup>'s Generally Accepted Recordkeeping Principles [1] define a global standard that identifies the criticalities and a high-level framework of good practices for information governance. They are a common set of principles that describe the conditions under which business records and related information should be maintained. They are:

**Accountability:** there should be a person that is responsible and accountable for all the process;

**Transparency:** all the information should be documented in an open and verifiable manner;

**Integrity:** the information assets should be as authentic and reliable as possible;

**Protection:** no unauthorized parties should be able to access private information;

**Compliance:** laws and policies should be kept into consideration;

**Availability:** information should be efficiently and accurately retrieved;

**Retention:** information should remain accessible for a period of time depending on legal, regulatory, fiscal, operational, and historical requirements;

**Disposition:** it should be possible to erase all the information that is no longer needed.

The ISO Standard 15489-1:2016 Information and documentation — Records management [4] defines concepts and principles for the creation, acquisition and management of records. Section 7.2 *Characteristics of a record* lists the following:

**Authenticity:** records must be created and maintained in such a way that creators are authorized and identified, and that records are protected against unauthorized addition, deletion, alteration, use and concealment

**Reliability:** the content of a record should be accurate, and its creator should be worth of trust

---

<sup>8</sup> ARMA International is a not-for-profit professional association and a global authority on governing information as a strategic asset

**Integrity:** a record should be complete and protected against unauthorized alteration. Every alteration should be documented and traceable

**Usability:** a useable record is one that can be located, retrieved, presented and interpreted.

## CHALLENGES AND PERSPECTIVES

Even if those guidelines have been defined having in mind mainly the management of archives in business activities, some of them are very generic and can be referred to any digital archive. An overview of the most relevant features associated to a digital archive can be found in [11]. Here, we focus only on reliability, authenticity, and availability over time. Reliability has a lot to do with the trustworthiness of the creator of a record, and their competence to capture the truth, so it is necessary to define an effective method of authentication that avoids problems of plagiarism. Authenticity regards the identity of a record, and its integrity. Finally, availability over time means providing long-term access to the information. Long-term refers to a period of time that is long enough to survive to changing technologies, such as new media and data formats.

Ensuring a correct access to digital archives includes the following challenges:

### **Digital Preservation**

Digital preservation guarantees that digital objects are available and usable over the time [2]. A digital archive should take care of storage media instability and deterioration, which could lead to data loss, and technology obsolescence and incompatibility, which may happen both at the hardware and software level.

### **Decentralization**

A digital archive is often owned by a centralized organization, such as a government or a library, which stores all digital objects in a local database and then publishes them on the Web. This means that the archive leaves as long as the centralized organization owning it. Decentralization should guarantee that the ownership of a digital archive is not associated to any particular organization. Digital objects should belong to the human heritage and they should be accessible and available everywhere and every time, despite the organization hosting the corresponding physical copies.

### **Security Issues**

Security should guarantee the integrity of digital objects contained in the archive, i.e. digital objects should not be modified by not authorized entities. In addition, a secure digital archive should prevent plagiarism, i.e. it should guarantee that only authorized digital objects are added to the archive.

The use of blockchain as main repository for digital archives could overcome all the described challenges. A blockchain is intrinsically distributed, thus all the information it stores is replicated on all the nodes of the network. This guarantees that stored objects are preserved over the time. In fact, if a failure occurs on a node of the network, information is not lost.

A blockchain guarantees also decentralization, in the sense that data are not hosted by a single central authority.

Finally, a blockchain overcomes the described security issues because of its intrinsic nature, based on cryptography.

It is worth mentioning, though, that identity authentication is performed by checking if a transaction is signed with a correct private key. In other words, identity is associated with key ownership, with no guarantees over the real identity of the owner of that key.

The topic of blockchain has been covered also in terms of accountability and transparency [5].

At the moment, these aspects are out of the scope of this paper.

## A POSSIBLE SCENARIO

Blockchain technology could be used to implement a platform for the registration of minor artworks. We refer to works that are artistically relevant but not as well-known as famous masterpieces, or belonging to the so-called minor arts, such as books and manuscripts, pottery, lacquerware, furniture, jewellery, or textiles. The aim would be to preserve artwork that, since it is often not well protected, can easily become subject to counterfeiting, be stolen or damaged by natural disasters. Examples of such works could be those kept in some small libraries or churches, or even in private households.

In details, the platform should guarantee the following aspects:

**Protection:** protect the digital description of the artwork in case of natural disasters and/or attacks (it is obviously impossible to protect the real work only with IT tools);

**Anti-counterfeiting:** protect artworks against forgery and allow correct identification of works in case of loss and subsequent discovery;

**Integrity:** make sure that the digital description of the artwork is not subject to unauthorized changes;

**Traceability:** trace all movements of individual artworks.

The previous objectives can be achieved through the implementation of a service based on blockchain. Protection would be achieved through the fact that the blockchain is replicated on different nodes. Anti-counterfeiting would be guaranteed by associating each work to a sort of digital identity card, containing all the information related to the work (including physical information). Finally, integrity and traceability would be intrinsically guaranteed by the immutability and timestamping properties of the blockchain. In fact, blockchain security assumptions guarantee that if at a certain time a piece of information has been added to a block that reached consensus, it will be impossible to alter that information without altering all the following blocks.

In this implementation, each minor artwork is associated to an Identity Card (IC), which contains all the metadata about the artwork: of course its author, the date and place of creation, as well as its current owner and location, and all the other relevant information that is dependent on the specific kind of art, along with administrative and technical metadata.

An IC could be implemented through a smart contract having three main methods: `insertArtwork(data)` and `updateArtwork(data)`, which modify the blockchain, and the function `getInformation(artwork)` that accesses data stored in the contract.

The collection of all artworks is stored into a permissioned blockchain, which is composed of different nodes, hosted by some authorized organizations, such as governments and representative institutions. The fact that different institutions have a copy of the blockchain helps reducing errors and incompatibilities in the registration and management of the artworks.

Two types of users can access the blockchain of artworks: *artwork owners* and *artwork checkers*. Artwork owners host artworks (e.g. museums) and are responsible of inserting the IC of new artworks in the blockchain. In addition, artwork owners update ICs, whenever a change occurs, such as a temporary movement of the artwork for an exhibit or a restoration. Artwork checkers are responsible of verifying that artworks are not counterfeit or stolen. Examples of artwork checkers could be police departments or art experts.



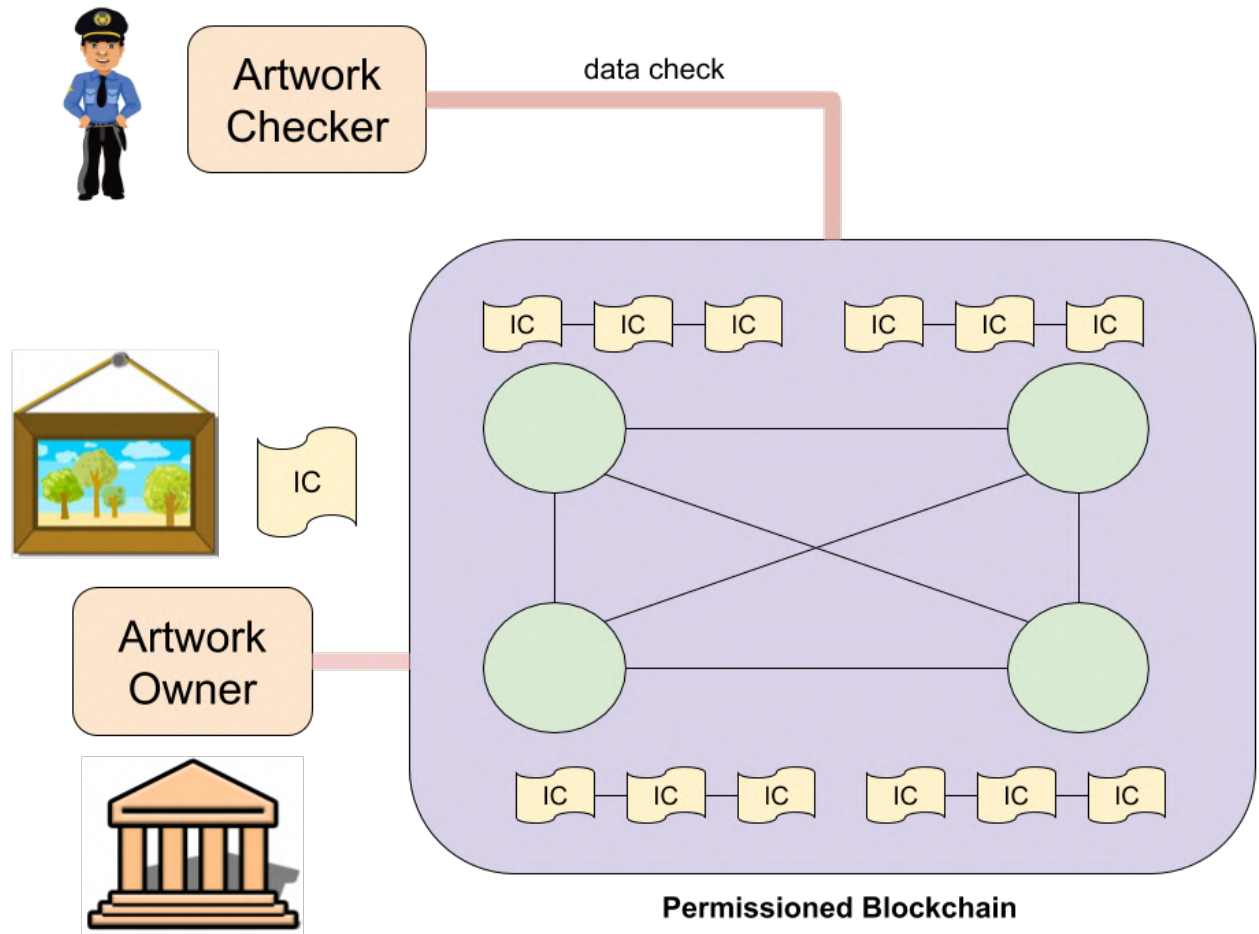


Figure 1: A possible implementation of the service for the registration of minor artworks.

## REFERENCES

- [1] ARMAcx International. Generally Accepted Recordkeeping Principles. 2017.
- [74] Borghoff, U. M., Rödiger, P., Scheffczyk, J., and Schmitz, L. *Long-Term Preservation of Digital Documents: Principles and Practices*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [75] Cachin, C. *Architecture of the hyperledger blockchain fabric*. In Workshop on Distributed Cryptocurrencies and Consensus Ledgers. 2016.
- [76] ISO. ISO 15489-1/2: 2016- Information and documentation - Records management. 2016.
- [77] Lemieux, Victoria L. *The Future of Archives as Networked, Decentralized, Autonomous and Global*. In Caroline Brown, Ed. *Archival Futures*. London: Facet Publishing (forthcoming).
- [78] Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system.
- [79] Schwartz, D., Youngs, N., Britto, A., et al. *The ripple protocol consensus algorithm*. Ripple Labs Inc WhitePaper. 2014.
- [80] Swan, M. *Blockchain: Blueprint for a new economy*. O'Reilly Media, Inc.. 2015.
- [81] Wood, G. Ethereum: A secure decentralised generalized transaction ledger. Ethereum Project Yellow Paper. 2014.
- [82] Zeilinger, M. Digital art as 'monetised graphics': Enforcing intellectual property on the blockchain. *Philosophy & Technology*, 31(1), 15-41. 2018.

[83] Zheng, Z., Xie, S., Dai, H.-N., and Wang, H. *Blockchain challenges and opportunities: A survey*. Work Pap. 2016.

# L'Orientale Foundation: un programma per l'introduzione delle nuove tecnologie nei curricula umanistici

Johanna Monti<sup>1</sup>, Valeria Caruso<sup>2</sup>

<sup>1</sup>Università degli studi di Napoli L'Orientale, Italia – [jmonti@unior.it](mailto:jmonti@unior.it)

<sup>2</sup>Università degli studi di Napoli L'Orientale, Italia – [vcaruso@unior.it](mailto:vcaruso@unior.it)

## ABSTRACT

Il connubio di saperi umanistici, tecnologia e competenze trasversali (le cosiddette *soft skill*), ovvero creatività, valori, pensiero creativo e indipendente, capacità di lavorare in gruppo, rappresenta un elemento ormai imprescindibile per il futuro lavorativo dei laureati in discipline umanistiche. L'Orientale ha sviluppato, in collaborazione con la Apple *L'Orientale iOS Foundation Program*, un progetto formativo di quattro settimane per l'acquisizione sia di competenze digitali, in particolar modo per lo sviluppo di App, sia di competenze trasversali come l'addestramento al team work, alle strategie di problem-solving o alla comunicazione pubblica delle proprie idee e progetti.

Il corso si avvale di un approccio didattico innovativo (*Challenge Based Learning*), basato sugli assunti della pedagogia costruttivista secondo cui le conoscenze individuali sono il frutto di un processo di costruzione attiva del sapere, non il risultato di una trasmissione di informazioni. Gli obiettivi formativi vengono infatti raggiunti mediante un processo di ricerca, discussione e lavoro di gruppo attraverso cui gli studenti progettano ed implementano una App che sia di supporto alla soluzione di un problema concreto da loro avvertito come 'sfida' ('challenge' in inglese). Il *Challenge Based Learning* inizia infatti dall'identificazione di una questione che stia a cuore ai discenti e si sviluppa attraverso una serie di attività di esplorazione e ricerca che porteranno alla realizzazione di una App.

All'interno di questo *framework* didattico, le lezioni dedicate alle conoscenze informatiche si configurano come i tasselli di un ciclo di apprendimento interattivo, in cui i linguaggi di programmazione (nel nostro caso Swift 4.0) e gli ambienti di sviluppo software (come Xcode) diventano strumenti necessari alla realizzazione di un tool (la App).

Il contributo discute i risultati ottenuti durante i primi tre corsi de *L'Orientale iOS Foundation Program* sia in relazione ai tipi di App realizzate e al loro livello di implementazione, sia sintetizzando i dati dei questionari di fine corso somministrati agli studenti.

## PAROLE CHIAVE

Coding; digital humanities; IT Humanists; App development; Humanities Apps.

## INTRODUZIONE

Il World Economic Forum nel 2016 [1] ha messo in evidenza come nella quarta rivoluzione industriale (Industria 4.0) sia necessario sviluppare competenze trasversali (le cosiddette *soft skill*), quali ad esempio la capacità di lavorare in gruppo, risolvere problemi complessi, pensare con flessibilità, unitamente a quelle collegate all'utilizzo delle tecnologie digitali. Sulla scia di questi orientamenti, l'Università degli studi di Napoli "L'Orientale" è impegnata dal 2016 nella realizzazione de *L'Orientale iOS Foundation Program*, in cui poter acquisire le competenze richieste da un mercato del lavoro in continua evoluzione. Il programma si sviluppa in quattro settimane di corso intensivo, con sette ore di attività in aula al giorno dedicate al design e sviluppo di App, all'apprendimento del linguaggio di programmazione

del mondo iOS (*Swift Programming Language*) e ai *soft skill*. Al termine del corso, i discenti presentano il prototipo di una loro App realizzata all'interno di gruppi di lavoro.

Le lezioni di programmazione sono impartite da un informatico, mentre un gruppo di istruttori guida la classe nello sviluppo delle abilità di *soft skill* e del *concept* dell'App, cercando di focalizzare la loro attenzione su strumenti che siano di supporto al mondo umanistico. Mentre i docenti di *coding* sono dei liberi professionisti, gli istruttori sono docenti e ricercatori de *L'Orientale* selezionati per il *Foundation Program* che, al pari dei colleghi programmatori, hanno ricevuto e continuano a ricevere una formazione specifica da parte di Apple.

Illustrando l'approccio didattico adottato, ovvero il *Challenge Based Learning*, il contributo discute i risultati ottenuti durante i primi tre corsi de *L'Orientale iOS Foundation Program*, sottolineando come alcune innovazioni metodologiche introdotte nell'ultima edizione abbiano contribuito al miglioramento del programma in relazione a (i) i tipi di App realizzate, (ii) al loro livello di implementazione e (iii) al livello di competenze raggiunto dai corsisti. Vengono presentati a tal fine i dati dei questionari di fine corso somministrati agli studenti e alcuni prototipi di App da loro sviluppati.

### **La metodologia didattica del corso: il Challenge Based Learning**

Il *Challenge Based Learning* (CBL) è un approccio didattico collaborativo basato su delle sfide che vengono identificate e risolte dai partecipanti stessi, sulla base di proposte solutive che devono essere implementate e verificate. Il processo, scandito da fasi ben precise, consente ai partecipanti di acquisire i) conoscenze approfondite e ii) competenze richieste dall'attuale mercato del lavoro.

Nato come approccio didattico dal progetto "Apple Classrooms of Tomorrow—Today" (ACOT2) nel 2008 per identificare i principi essenziali dell'ambiente di apprendimento nel XXI secolo [2, 3, 4], il *Challenge Based Learning* è basato sull'apprendimento esperienziale e, in particolare, sull'esplorazione della realtà e su una riflessione incentrata sul percorso di apprendimento stesso, ispirandosi a idee innovative nel campo della tecnologia, dei media, del mondo del lavoro e della società attuale [5]. Questo metodo didattico fornisce un quadro di riferimento per imparare identificando, comprendendo e risolvendo sfide autentiche, nonché intraprendendo azioni efficaci per raggiungere un risultato. Le linee-guida teoriche della CBL affondano nel modello pedagogico costruttivista di Seymour Papert [6] che, ispirandosi a Jean Piaget, intende l'apprendimento come "costruzione di strutture conoscitive" realizzabile con efficacia se il discente è coinvolto attivamente nel processo di edificazione.

L'approccio, dopo una sperimentazione avvenuta nel corso dell'ultimo decennio principalmente nelle scuole americane [5], approda alla sua forma attuale, ovvero un processo scandito in tre fasi (Fig. 1):

- Engage (coinvolgimento) – attraverso un processo di identificazione delle cosiddette domande fondamentali (Essential Questioning), i partecipanti identificano una sfida concreta da affrontare a partire da un'idea astratta (Big Idea).
- Investigate (ricerca) – i partecipanti pianificano e partecipano ad un processo in cui cercano le soluzioni alla sfida, in primo luogo attraverso la ricerca delle risorse necessarie (ad esempio attraverso l'uso di Internet, riferimenti bibliografici, ricerche sul campo, ecc.) poi attraverso la pianificazione delle azioni necessarie per affrontare la sfida individuata (questionari, interviste, sondaggi, etc.).
- Act (azione) – vengono sviluppate delle soluzioni sulla base di un pubblico reale, che vengono successivamente valutate in base ai risultati attesi.

L'esperienza del corso, ed in particolare di una modalità didattica collaborativa ed esperienziale, è nuova e molto apprezzata dagli studenti, ma richiede anche un notevole sforzo da parte loro, che devono abbandonare l'atteggiamento passivo con cui di solito seguono le lezioni frontali nelle aule universitarie. Nel *Challenge Based Learning*, al contrario, è necessario diventare co-costruttori del proprio percorso formativo. Per questo motivo, è necessario disporre di ambienti progettati specificamente per la formazione dei gruppi e per favorire la collaborazione tra i partecipanti: l'aula è arredata con tavoli tondi e dispone di una rete *wifi* a banda larga, di portatili di ultima generazione e di quattro maxischermi visibili da qualsiasi posizione.

## **I corsisti**

La partecipazione a *L'Orientale iOS Foundation Program* è gratuita ma riservata agli iscritti ad un corso di laurea o di dottorato dell'Ateneo, previa selezione. Per iscriversi al programma, gli studenti devono superare una prova scritta e inviare un video motivazionale attraverso cui far emergere la propria creatività. Il test scritto consiste in un quiz online a risposte multiple e viene somministrato attraverso la piattaforma elearning de *L'Orientale* (gestita con Moodle), piattaforma che peraltro ospita tutti i contenuti e materiali didattici del corso. Oltre ai quesiti di logica, il test verte sui linguaggi di programmazione, in particolare quelli ad oggetti (*OOP – Object Oriented Programming*), e sulle caratteristiche e le risorse del mondo iOS.

La partecipazione al programma consente di acquisire crediti per tirocini e/o altre attività formative, in base al tipo di riconoscimento che è stato stabilito dai diversi corsi di laurea.

## **Le lezioni**

Nelle quattro settimane di corso i contenuti didattici sono distribuiti in modo da concentrare gli elementi di maggiore novità e complessità tra la metà della prima e della terza settimana, dopo una introduzione generale che guida i discenti nello sviluppo dei *soft skill* e li avvicina all'apprendimento basato sulle sfide della CBL.

Durante la seconda settimana, i partecipanti approfondiscono le diverse fasi della CBL (*Engage, Investigate, Act*), cominciando a riflettere su questioni della vita reale per poi scegliere un problema specifico che diventa la sfida (*challenge*) a cui rispondere con una App da loro implementata. La CBL li guida nell'analisi delle caratteristiche del problema scelto e nel reperimento delle informazioni e risorse utili a risolverlo. Impegnati in questi processi, i discenti sviluppano anche competenze trasversali come la capacità di gestire il tempo, il lavoro di gruppo e i conflitti; ma anche l'abilità di argomentare e difendere le proprie opinioni. Ai docenti-istruttori spetta il compito d'illustrare il modello esplorativo della CBL e di monitorarne l'apprendimento, osservando le dinamiche interattive dei gruppi, l'avanzamento del lavoro e rispondendo a dubbi e perplessità sull'applicazione del metodo stesso. Il successo dipende molto da quanto le sfide scelte siano coerenti con le passioni e le competenze dei partecipanti e gli istruttori sono impegnati a fare in modo che ciò possa realizzarsi in tutti i gruppi di lavoro.

Il momento cruciale per le attività del corso si realizza verso la fine della prima settimana quando i) si propongono delle attività per formare i gruppi che lavoreranno insieme fino alla fine, ii) i gruppi decidono quale sarà la sfida finale che intendono perseguire.

Nella prima settimana vengono invece fornite informazioni generali sul corso e cominciano le lezioni di programmazione, ma molte delle attività sono finalizzate alla socializzazione e alla formazione dei gruppi. È necessario, in questa fase, che i discenti acquisiscano una maggiore consapevolezza di se stessi e di come la loro personalità e le loro abilità possano contribuire al *team work*. Alcuni *task* vertono proprio sull'autopresentazione e sulla definizione delle caratteristiche psicologiche di ciascuno, tratteggiate a partire dall'inventario delle personalità

del *Myers-Briggs Type Indicator* (MBTI). Ad ogni tipo di profilo viene assegnato un ruolo in un team di lavoro provvisorio che risolverà una *nano-challenge* per sperimentare la didattica della CBL e l'*engagement* tipico delle sue sfide.

### Implementazioni didattiche de L'Orientale iOS Foundation Program

A fronte di una strutturazione tanto serrata delle fasi del processo di apprendimento, il modello della CBL è stato però concepito per essere flessibile e personalizzabile: “A flexible and customizable framework that can be implemented as a guiding pedagogy or integrated with other approaches to learning” [2]. Ad esempio, durante l'ultimo corso de *L'Orientale iOS Foundation Program* il paradigma è stato implementato per focalizzare l'attenzione dei discenti sul mondo delle App umanistiche, in modo da orientare le attività di ideazione e programmazione verso questo ambito. Sono state proposte in tal senso attività di esplorazione (o *nano-challenges*), realizzate sia dai docenti sia dai discenti, delle risorse disponibili sul mercato che si occupano di problematiche del mondo umanistico, come App e servizi digitali per la traduzione, per l'archeologia o per l'apprendimento delle lingue. Inoltre, nell'ultimo corso realizzato, sono stati indicati dei macro-temi da cui partire per approdare alla *challenge* finale, alla quale i discenti lavorano durante tutto il corso per realizzare il prototipo di una loro App. La scelta è ricaduta sui seguenti temi: ambiente, disabilità, educazione, viaggi e percorsi alternativi.

L'introduzione di questi indicatori iniziali ha consentito di incrementare il coinvolgimento dei discenti nel progetto della loro App, realizzando *concept* di ambito umanistico e prototipi ben implementati, quasi pronti per essere pubblicati. Analogamente, nei questionari di fine corso sono stati riscontrati incrementi nelle competenze che i discenti ritengono di aver sviluppato, rivelando un livello di confidenza maggiore con i contenuti del corso. Si veda in tal senso la tabella 1 sottostante, che sintetizza in maniera comparativa le risposte date alle domande del *final survey* dai corsisti dei primi due corsi e quelli del terzo. Come si evince, il terzo corso ha contribuito in maniera più incisiva ad implementare gli *skill* dei discenti.

Please rate how confident you are to be able to do the following as of now	Positive answers (course 1 e 2)	Positive answers (course 3)
Design an user experience (UX)	86%	100%
Make a prototype	91%	100%
Collaborate in a team to create an app	100%	100%
Manage an app project	97%	100%
Learn more about iOS development by myself	100%	100%
Pitching your app/software idea to a potential investor	91%	96%
Publishing an app to the App store	79%	83%
Work as a professional developer	53%	79%
Start your own mobile business	62%	71%

Tabella 1. Risultati comparativi, collezionati con survey di fine corso, sul livello di confidenza sviluppato dai discenti per diverse abilità durante i primi e il terzo corso de L'Orientale iOS Foundation Program.

La valutazione degli obiettivi formativi attesi per il corso si basa invece esclusivamente sul *concept* delle App e sul loro livello di implementazione. I corsisti ricevono un attestato di partecipazione senza voti, in ragione anche del modello pedagogico della CBL, in cui

l'*engagement*, ovvero la partecipazione e l'interesse individuale, è l'unico parametro che possa stabilire il successo della didattica. Le esperienze maturate in questi due anni di formazione rivolta agli umanisti sembrano peraltro confermare la bontà di questo assunto, dal momento che alcuni dei corsisti de *L'Orientale iOS Foundation Program* (circa 10 su 60) hanno, per il momento, abbandonato le Humanities e si sono diretti verso percorsi di formazione, in alcuni casi anche di lavoro, nel campo dell'informatica. Si spera che questi ex-corsisti riescano, nel tempo, a coniugare i due diversi tipi di *skill* acquisiti e diventino compiutamente degli *IT-Humanists*.

Per il momento, intendiamo sottolineare la validità delle implementazioni introdotte durante i corsi de *L'Orientale iOS Foundation Program* per favorire l'apprendimento degli *skill* di programmazione necessari a sviluppare una App con Xcode (l'ambiente di sviluppo per iOS). Il focus sui software per gestire problematiche di ambito umanistico ha di fatto incrementato il coinvolgimento degli studenti universitari di queste discipline e si è tradotto in risultati migliori in termini di (i) *concept* di App, che risultano più innovative (come nel caso di *Deathaly* [8, 9]), (ii) livello di implementazione del software sviluppato dai discenti, (iii) livelli di confidenza da parte degli studenti sulle competenze che il corso dovrebbe aiutare a sviluppare.

## BIBLIOGRAFIA

- [1] The Global Competitiveness Report 2015–2016  
[[http://www3.weforum.org/docs/gcr/2015-2016/Global\\_Competitiveness\\_Report\\_2015-2016.pdf](http://www3.weforum.org/docs/gcr/2015-2016/Global_Competitiveness_Report_2015-2016.pdf)]
- [2] Nichols, M., Cator, K., Torres, M., Henderson, D. 2016. *Challenge Based Learner User Guide*. Redwood City, CA: Digital Promise.
- [3] Apple 2008. *Apple Classrooms of Tomorrow - Today*.  
[[http://ali.apple.com/acot2/global/files/ACOT2\\_Background.pdf](http://ali.apple.com/acot2/global/files/ACOT2_Background.pdf)]
- [4] Apple Education 2010. *Challenge Based Learning Classroom Guide*.  
[[http://cbl.digitalpromise.org/wp-content/uploads/sites/7/2016/08/CBL\\_Classroom\\_Guide2010.pdf](http://cbl.digitalpromise.org/wp-content/uploads/sites/7/2016/08/CBL_Classroom_Guide2010.pdf)]
- [5] Sandholtz, J.H., Ringstaff, C., Dwyer, D.C. 1997. *Teaching with Technology: Creating Student-Centered Classrooms*. New York: Teachers College.
- [6] Papert, S. Harel, I. (eds). 1991. *Constructionism: research reports and essays 1985-1990 by the Epistemology and Learning Research Group*. The Media Lab, Massachusetts Institute of Technology, Norwood, NJ: Ablex Pub. Corp.
- [7] Briggs Myers, I. 1998. *MBTI Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator*. Palo Alto: Consulting Psychologists Press.
- [8] Festa, N. 2018. “La Napoli di ‘questi’ fantasmi. Un app racconta storie (e miti)”. *Corriere del Mezzogiorno*, 26 settembre.
- [9] Capone, M. 2018. Spunta la app che acchiappa i fantasmi”, *Il Mattino* – Napoli.

# Towards a Computational Stylistics of Book Reviews

Massimo Salgaro<sup>1</sup>, Simone Rebora<sup>2</sup>

Università di Verona<sup>1</sup>, Georg-August-Universität Göttingen<sup>2</sup>  
*massimo.salgaro@univr.it, simone.rebora@phil.uni-goettingen.de*

## ABSTRACT

With this paper, we set up the groundwork for the computational study of the style of literary criticism. Given a corpus of Italian book reviews published in a social reading platform, a paper magazine, and three scientific journals, we evaluate to what extent stylometry can be used to distinguish professional from non-professional reviewers. In the first part, we discuss the potential (and limitations) in the use of most frequent words (MFW) and distance measures: while even lower selections of MFW score positive results, the average length of book reviews does not seem sufficient for reaching an acceptable classification quality. In the second part, we introduce a machine learning framework that combines multiple dictionary-based measurements (aimed at evaluating the presence of a hyper-specialized lexicon, or of emotional aesthetic response and mental imagery) to train a Support Vector Machines (SVM) classifier. Results show how the combination of traditional stylometric methods and dictionary-based measurements helps improving classification accuracy, even if the most effective feature is still the statistical distribution of function words. We discuss the theoretical implications of our discoveries (that seem to contradict—or at least scale back—traditional definitions of literary criticism) and we propose possible future developments of the project, whose final goal will be that of better understanding how the Internet is changing the way books are read and evaluated.

## KEYWORDS

Literary criticism, Book reviews, Stylometry, Word class analysis, Machine learning.

## INTRODUCTION

Digital reading and the Internet have not only changed our way of reading literature but also the way to acquire texts, to buy them and to discuss about them. While in the past readers found information on books in libraries, newspapers, journals, TV-programs, and recommendations by friends or family, during the past decade, the Internet has gained in importance as an additional source of information (Šesek and Pušnik 2014; Schreier 2010; Rohmer 2001; Hargittai 2000). Contemporary readers can find book reviews in author's websites, fan sites, publisher's and bookseller's websites (e.g. Amazon.com). These new hosts of literary reviews created new forms of literary criticism. Reviewers in newspapers and journals act as experts, guiding the public and pinpointing the literary status and reputation of a book. Nowadays, with the social media and social reading platforms like Goodreads, "readers are increasingly taking over as critics" (Schreier 2010: 309), informing the reading selection of their peers. While the traditional professional critic is detached from the book he/she is reviewing (Eco 1979), these spontaneous, non-professional reviewers are guided by emotional/cognitive involvement, entertainment and even identification with the book characters (Schreier 2010: 312). So far, studies on these new forms of literary criticism and literary reviews are still lacking (among the few exceptions, Murray 2018; Faggiolani and Vivarelli 2016; Cordon-García et al. 2013), while their importance in guiding the book selection of other readers is beyond all questions. What is also lacking is a linguistic description of the stylistic cues that distinguish the reviews written by non-professional



readers—while only a few attempts have been made to study the technical terminology of professional readers (Timm 1992; Gardt 1998; Beck et al. 2007).

### THE POTENTIAL (AND LIMITATIONS) OF STYLOMETRIC ANALYSIS

In order to verify to what extent stylometric methods can be used to classify the reviews written by professional and non-professional readers, we adopted a corpus of Italian book reviews (Salgaro and Reborá 2018). The corpus was divided into three subsets: reviews published on social reading platforms (source: [aNobii](#)), in paper magazines ([Il Sole 24 Ore](#)), and in scientific journals ([Between](#), [Osservatorio critico della germanistica](#), and [OBLIO](#)). All sub-corpora had an approximate size of 650,000 tokens. All details are described in Table 1.

	Social Reading	Paper Magazines	Scientific Journals
Source	aNobii	Sole 24 Ore	Between – Osservatorio critico della germanistica – OBLIO
Publication dates	2007-2017	2010-2011	1998-2016
Total number of tokens	646964	704242	655192
Length of shortest review (tokens)	1	32	234
Length of longest review (tokens)	2229	1840	3195
Mean length of reviews (tokens)	125.3174	526.5012	1197.275
Standard deviation of review length	169.9646	350.4949	459.4581

Table 1. Corpus composition

As demonstrated by (Eder 2015), the first element to influence the quality of a stylometric classification is text length. As Eder suggested a minimum of 5,000 words, we ran a series of preliminary tests with the [Stylo](#) R package (Eder et al. 2016) on artificially-composed 5,000-word-long text chunks (i.e. shorter reviews from the same subsets were concatenated). Results show how Cosine Delta distance, based on just 50 most frequent words (MFW), was able to almost perfectly separate the three subgroups (see Fig. 1).

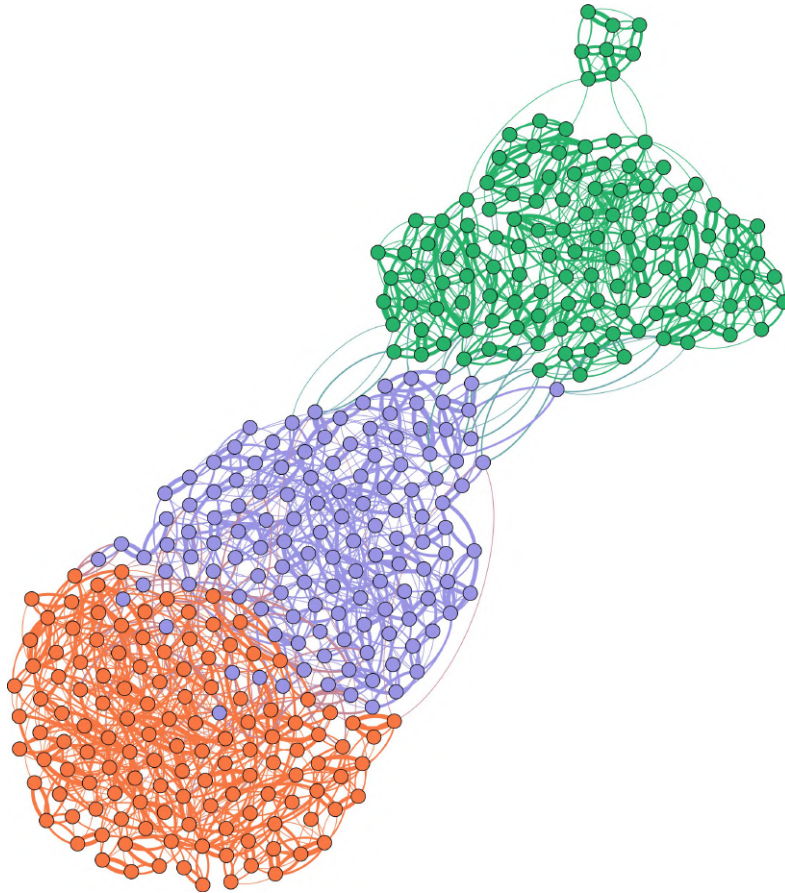


Figure 1. Network graph of the corpus (Cosine Delta, 50 MFW, ForceAtlas2 in [Gephi](#)). In green: aNobii; in violet: Sole 24 Ore; in red: Between, Osservatorio, and OBLIO.

Considering the high variance of text length in our corpus (mean = 259 words; standard deviation = 363 words), we set up an experiment with the [PyDelta](#) Python library. First, we generated a series of sub-corpora composed by text chunks of the same length, varying between 50 and 5,000 words. For each configuration, we then evaluated clustering quality through the adjusted Rand index (see Everitt et al. 2011: 264f.). Figure 2 shows how Cosine Delta distance with 2,000 MFW is the best performing classifier (confirming the results already obtained by Evert et al. 2017). However, it can be noted how 200 MFW (i.e., mainly function words) reach an even better efficiency with longer text chunks. As for text length, clustering quality is quite poor below 1,000 words, while a plateau is reached at about 3,000 words.



Figure 2. Clustering quality per slice length and MFW used (distance: Cosine Delta).

## BEYOND MOST FREQUENT WORDS: A MACHINE LEARNING FRAMEWORK TO CLASSIFY BOOK REVIEWS

The results obtained with traditional stylometric methods are promising, but still not applicable to a real-case scenario, where reviews (especially when published on social reading platforms) can be just a few hundreds of words long. To improve the results for shorter chunks, we developed a framework for a machine learning classifier, by operationalizing a series of traditional definitions of literary criticism:

- to identify the technical skills of experienced readers (Eco 1979), an extensive lexicon of literary criticism (Beck et al. 2007) was translated into Italian;
- to identify the direct involvement of non-professional reviewers (Schreier 2010), selections of terms related to (a) mental imagery and (b) emotional aesthetic response were extracted from questionnaires and tools in empirical aesthetics, such as the “Story World Absorption Scale” (Kuijpers et al. 2014), the “enactment-imagery” scale (Fialho et al. in press), and (Knoop et al. 2016). The list was translated into Italian and expanded through the [fastText Italian](#) word-embedding model (Joulin et al. 2017);
- to complete the selection of features, (a) “Social Words”, (b) “Affect Words”, and (c) “Biological Processes” were identified through the [LIWC](#) Italian dictionary (Agosti and Rellini 2007). Even if promising in itself, the “Cognitive Processes” feature was excluded because of the contradicting results it produced in previous experiments (cf. Salgaro and Rebora 2018).

These resources were used to measure emotional and cognitive involvement with the reviewed texts. Each review was thus described by a total of six “dimensions”, which were

combined with the results of the stylometric analysis (Cosine Delta, 2,000 MFW) and used to train a Support Vector Machines (SVM) classifier.

For a corpus composed by 250-word-long chunks (the mean length of the texts in our corpus), the sole stylometric analysis reached an attribution accuracy of 82.9%, while the SVM classifier scored 87.4%. A slight but promising improvement, if we consider the simplicity of the framework—that can and should be refined further. Note also that, by increasing text length to 500 words, accuracy increases to 90.1% for Cosine Delta and to 93.2% for SVM.

## CONCLUSIONS AND FUTURE WORK

Among the features that can be used to strengthen the framework, the tendency of professional reviewers towards vagueness (Gardt 1998: 1356) can be operationalized through a resource like Wordnet—as already done by (Piper 2016) to distinguish the concreteness of fiction from the abstractness of non-fiction. The higher use of rhetorical figures (Gardt 1998: 1357) can be measured through approaches such as (Dubremetz and Nivre 2015) or (Mitrović et al. 2017). The dimension of complexity can also be added through simple scores like type-token ratio, the Flesch–Kincaid index, and many others. However, it should be noted how the results presented in this paper already contribute significantly to a better understanding of the “style of criticism.” Differently from what theoreticians have supposed so far, it seems that most of the distinction is played in the frequency of use of a few most frequent (function) words, and not in the higher-level features that we have operationalized here. Note for example how the SVM classifier, when working on the sole dictionary-based measurements for 500-word-long chunks, reaches just a 75.7% of accuracy. Future work will focus on identifying which one, among these high-level features, plays the most relevant—or irrelevant—role in the classification.

Among the possible future goals of this research, is also an analysis of how the very concept of literary value can be influenced by the growing phenomenon of social reading. In fact, the aim of book reviews is to attribute literary value to a text. Following Heydebrand and Winko, the attribution of literary value denotes “a complex social act by which a subject attributes value to a literary text, in a concrete situation and on the basis of a certain standard of value and certain categorizing assumptions” (Heydebrand and Winko 2008: 226). Consequently, a literary text is not intrinsically valuable, it only acquires an attributive value in relation to standards of value. For example, for professional reviewers, a “good” book should be “complex” or “rhetorically elaborated” to meet their expectations (values), whilst for non-professional readers, a “good” book can be a “suspenseful” love story or a “thrilling” detective novel. As Heydebrand and Winko point out, “literary evaluation is by no means limited by professional judgement on literary texts” (Heydebrand and Winko 2008: 225), thus implicitly including all potential reviewers. It takes place in a complex social system and plays a role in the production, distribution, and reception of literature. As recent experiments have shown, the “literary value” of a text changes depending on the material support the text is published on (Salgaro et al. 2018). Texts on paper and texts on screen seem to have still a different social prestige and this may also be true for literary reviews.

With this paper, we hope to have cast the groundwork for a research that might fruitfully combine computational methods and literary theory to study the “style of criticism” of professional and non-professional readers.

## NOTE

All scripts and resources used for the analysis are available here: [https://github.com/SimoneRebora/style\\_of\\_criticism](https://github.com/SimoneRebora/style_of_criticism).

## REFERENCES

- [1] Agosti, A. and A. Rellini. 2007. "The Italian LIWC Dictionary." Austin, TX: LIWC.net.
- [84] Beck, Rudolf, Hildegard Kuester, and Martin Kuester. 2007. *Basislexikon anglistische Literaturwissenschaft*. Paderborn: Fink.
- [85] Cordón-García, José-Antonio, Julio Alonso-Arévalo, Raquel Gómez-Díaz, and Daniel Linder. 2013. *Social Reading*. Oxford: Chandos.
- [86] Dubremetz, Marie, and Joakim Nivre. 2015. "Rhetorical Figure Detection: The Case of Chiasmus." In *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, pp. 23–31.
- [87] Eco, Umberto. 1979. *Lector in fabula: la cooperazione interpretativa nei testi narrativi*. Milano: Bompiani.
- [88] Eder, Maciej. 2015. "Does Size Matter? Authorship Attribution, Small Samples, Big Problem." *Digital Scholarship in the Humanities* 30 (2): 167–82.  
<https://doi.org/10.1093/llc/fqt066>.
- [89] Eder, Maciej, Mike Kestemont, and Jan Rybicki. 2016. "Stylometry with R: A Package for Computational Text Analysis." *The R Journal* 8 (1): 1–15.
- [90] Everitt, Brian S, Sabine Landau, Morven Leese, and Daniel Stahl. 2010. *Cluster Analysis*. Chichester, West Sussex: Wiley.
- [91] Evert, Stefan, Thomas Proisl, Fotis Jannidis, Isabella Reger, Steffen Pielström, Christof Schöch, and Thorsten Vitt. 2017. "Understanding and Explaining Delta Measures for Authorship Attribution." *Digital Scholarship in the Humanities* 32 (suppl. 2): ii4–ii16.  
<https://doi.org/10.1093/llc/fqx023>.
- [92] Faggiolani, Chiara, and Maurizio Vivarelli, eds. 2016. *Le Reti Della Lettura: Tracce, Modelli, Pratiche Del Social Reading*. Milano: Editrice Bibliografica.
- [93] Fialho, Olivia, Hans Hoeken, and Frank Hakemulder. in press. "Literary Imagination and Changing Perceptions of Self and Others: an Explanatory Model of Transformative Reading."
- [94] Gardt, Andreas. 1998. "Die Fachsprache Der Literaturwissenschaft Im 20. Jahrhundert." In *Fachsprachen*, edited by L. Hoffmann, H. Kalverkämper, and H.E. Wiegand, 1355–62. Berlin, New York: de Gruyter.
- [95] Hargittai, Eszter. 2000. "Open Portals or Closed Gates? Channeling Content on the World Wide Web." *Poetics* 27 (4): 233–253.
- [96] Heydebrand, Renate von and Simone Winko. 2008. "The Qualities of Literatures: A Concept of Literary Evaluation." In: *The Quality of Literature: Studies in Literary Evaluation*, edited by W. van Peer, Amsterdam & Philadelphia: John Benjamins Publishing Company, pp. 223-239
- [97] Joulin, Armand, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2017. "Bag of Tricks for Efficient Text Classification." In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, 427–431. Association for Computational Linguistics.
- [98] Knoop, Christine A., Valentin Wagner, Thomas Jacobsen, and Winfried Menninghaus. 2016. "Mapping the Aesthetic Space of Literature 'from Below.'" *Poetics* 56 (June): 35–49. <https://doi.org/10.1016/j.poetic.2016.02.001>.
- [99] Kuijpers, Moniek M., Frank Hakemulder, Ed S. Tan, and Miruna M. Doicaru. 2014. "Exploring Absorbing Reading Experiences. Developing and Validating a Self-Report Scale to Measure Story World Absorption." *Scientific Study of Literature* 4 (1): 89–122.
- [100] Mitrović, Jelena, Cliff O'Reilly, Miljana Mladenović, and Siegfried Handschuh. 2017. "Ontological Representations of Rhetorical Figures for Argument Mining." *Argument & Computation* 8 (3): 267–87.

- [101] Murray, Simone. 2018. *The Digital Literary Sphere: Reading, Writing, and Selling Books in the Internet Era*. Baltimore: Johns Hopkins University Press.
- [102] Piper, Andrew. 2016. "Fictionality." *Journal of Cultural Analytics*, December. <https://doi.org/10.22148/16.011>.
- [103] Rohmer, Ernst. 2001. "Linkliste und Leselust. Möglichkeiten der Literaturvermittlung im Internet." *Jahrbuch für Computerphilologie* 3: 43–54.
- [104] Salgaro, Massimo and Simone Rebora. 2018. "Measuring the 'Critical Distance'. A Corpus-Based Analysis of Italian Book Reviews." In *AIUCD2018 - Book of Abstracts*, edited by Daria Spampinato, 161–63. <https://doi.org/10.6092/unibo/amsacta/5997>.
- [105] Salgaro, Massimo, Pasqualina Sorrentino, Gerhard Lauer, Jana Lüdtkke, and Arthur Jacobs. 2018. "How to measure the social prestige of a Nobel Prize in Literature? Development of a scale assessing the literary value of a text", *TXT* 1: 138-148.
- [106] Schreier, Margrit. 2010. "'Is it Possible to Give a 6 out of 5 Stars?': Book Selection and Recommendation in the Internet Age." *Primerjalna književnost* 33 (2): 307–20.
- [107] Šesek, Lara, and Maruša Pušnik. 2014. "Reading Popular Literature and Digital Media: Reading Experience, Fandoms, and Social Networks." *Anthropological Notebooks* 20 (2): 103–26.
- [108] Timm, Christian. 1992. *Gibt es eine Fachsprache der Literaturwissenschaft?: fachtextlinguistische Untersuchungen an englischen Texten der Literaturgeschichtsschreibung*. Frankfurt am Main; New York: P. Lang.

# Lingue antiche e *Digital Humanities*: annotazione digitale dei testi tra ricerca e didattica

G. Mugelli, G. Re, A. Taddei  
Università di Pisa

## ABSTRACT

Questo articolo illustra le applicazioni per la didattica delle lingue antiche di un sistema di annotazione digitale di testi, nato dalla collaborazione tra il CoPhiLab del CNR di Pisa<sup>1</sup> e il Laboratorio di Antropologia del Mondo Antico<sup>2</sup>.

Alla base della cooperazione tra i due laboratori c'è l'esigenza di creare strumenti digitali che siano di supporto allo studio dei testi antichi da una prospettiva storico-antropologica.

## PAROLE CHIAVE

digital humanities, greco antico, latino, antropologia del mondo antico, annotazione, testi classici, euporia, ricerca, didattica, scuola, università.

## PER UNA DIDATTICA DIGITALE DELLE LINGUE ANTICHE

Da tempo si avverte l'esigenza di tentare forme di rinnovamento nella didattica delle lingue antiche, per restituire centralità all'esperienza liceale classica e all'apprendimento del latino e del greco. Il tema ha suscitato, negli ultimi anni, dibattiti e pubblicazioni che hanno avuto anche larga eco mediatica, che hanno avuto almeno il pregio di tornare a sollevare il problema della cosiddetta attualità dello studio delle lingue antiche, oggi<sup>3</sup>.

L'isolamento dello studio del latino e del greco dall'apprendimento di elementi di civiltà o dalle altre discipline rischia di costituire il principale veicolo della sua marginalizzazione: il liceo classico è una scuola valida non perché vi si insegnano *ancora* due lingue antiche, ma perché queste sono insegnate nella medesima scuola in cui vengono impartite altrettanto importanti discipline scientifiche.

Appare, dunque, necessario superare in modo definitivo il pregiudizio relativo alla incompatibilità tra IT e le cosiddette Scienze dell'Antichità, in modo da colmare il divario esistente tra la crescente familiarità degli studenti con gli strumenti informatici, da una parte, e la effettiva consapevolezza nell'uso di questi ultimi, dall'altra. Ancora di più questo appare valido per gli insegnanti i quali, per quanto di recente assunzione e formazione, non appartengono alla cosiddetta generazione Zeta. È nello spazio generazionale che separa i docenti dagli studenti che hanno concluso il liceo e hanno appena iniziato l'Università, che intende inserirsi il progetto qui descritto.

L'apprendimento delle lingue antiche necessita di una solida formazione di carattere storico e linguistico, ma è necessario non trascurare la componente semantica, finalizzata ad un apprendimento consapevole della componente morfologica e morfosintattica. Ad essere in gioco non è tanto una questione di apprendimento, quando piuttosto una funzione di potenziamento e personalizzazione delle conoscenze. Per questo motivo, da circa tre anni, si è scelto di sperimentare, nel *Laboratorio di lingua greca*, una forma di apprendimento che passa per un lavoro di annotazione digitale del testo.

---

<sup>1</sup> <http://cophilab.ilc.cnr.it:8080/CoPhiLabPortal/>

<sup>2</sup> <http://lama.fileli.unipi.it/>

<sup>3</sup> Vd. Roscalla 2016; Condello 2018.

## EUPORIA: ANNOTAZIONE USER-CENTRED PER LO STUDIO DEI TESTI ANTICHI

Euporia è un sistema di annotazione modellabile in base alle esigenze dell'utente, che si adatta a diversi campi di ricerca, dall'ecdotica all'ermeneutica<sup>4</sup>.

Il sistema di annotazione è concepito con un approccio *user-centered*: basandosi sull'adozione di Domain Specific Languages, il sistema è abbastanza flessibile da permettere di annotare fenomeni linguistici, morfologici, sintattici, o stilistici<sup>5</sup>, ma anche variantistica<sup>6</sup>. Il sistema può essere usato in esercizi di traduzione contrastiva<sup>7</sup> e per annotare diversi livelli di interpretazione del testo, proposti dall'annotatore stesso o rintracciabili nella letteratura secondaria.

Ogni DSL è strutturato secondo convenzioni che si avvicinano alle pratiche di annotazione degli esperti di un particolare campo. Le convenzioni stabilite dall'utente sono in un secondo momento convertite in sistemi di codifica standard tramite l'uso di un *parser*<sup>8</sup>: in particolare, le annotazioni sono esportate in TEI-XML, e i passaggi di testo citati sono convertiti in URI compatibili con i CTS<sup>9</sup>.

### EUPORIARAGT (RITUALS IN ANCIENT GREEK TRAGEDY)

Il progetto EuporiaRAGT è nato a supporto di una ricerca dottorale sulla forma e la funzione dei rituali nella tragedia attica<sup>10</sup>, nell'ambito della quale si osserva il funzionamento drammatico dei riti nel *corpus* delle tragedie superstiti, mettendo in rapporto la forma che il rituale drammatizzato assume con la realtà storica dell'Atene del V secolo a.C. Euporia è stato impiegato per un'annotazione tematica che incrocia dati sulle caratteristiche e le dinamiche del rito e dettagli della rappresentazione scenica (movimenti dei personaggi, azioni compiute dentro e fuori dalla scena).

---

<sup>4</sup> Il sistema di annotazione Euporia è sviluppato dal CoPhiLab dell'Istituto di Linguistica computazionale del CNR di Pisa, grazie alla collaborazione tra Federico Boschetti, Riccardo Del Gratta, Angelo Mario Del Grosso, Alessio Molinari, Luigi Bambaci, Andrea Bellandi e Fahad Khan.

<sup>5</sup> È attualmente in corso una annotazione di fenomeni retorici nel *corpus* ciceroniano, ad opera di Amedeo Alessandro Raschieri.

<sup>6</sup> Per quanto riguarda la *constitutio textus*, il sistema Euporia è stato adottato dal progetto di Luigi Bambaci, che sta studiando e annotando digitalmente la variantistica in più lingue del Qohelet. Sul linguaggio di annotazione e sulla possibilità di annotare varianti testuali o interpretative ci sia permesso di rinviare a Mugelli, Boschetti et al. 2016. Alcuni dei criteri di selezione delle varianti testuali si sono ispirati al progetto *MusisqueDeoque* (<http://www.mqdq.it>), vd. Mastandrea 2009.

<sup>7</sup> Il liceo classico "Gargallo" di Siracusa ha adottato, nell'a.s. 2017/2018, Euporia per un esercizio di traduzione contrastiva sui *Persiani* di Eschilo.

<sup>8</sup> ANTLR (<http://www.antlr.org>) permette di creare un *parser* che processa la sintassi delle annotazioni, espressa tramite una *context free grammar*.

<sup>9</sup> Canonical Text Services, vd. <http://www.homermultitext.org/hmt-doc/cite/index.html>.

<sup>10</sup> Gli aggiornamenti sul progetto EuporiaRAGT sono disponibili su <http://www.himeros.eu/euporiaRAGT/> e sul sito del [LAMA](#).



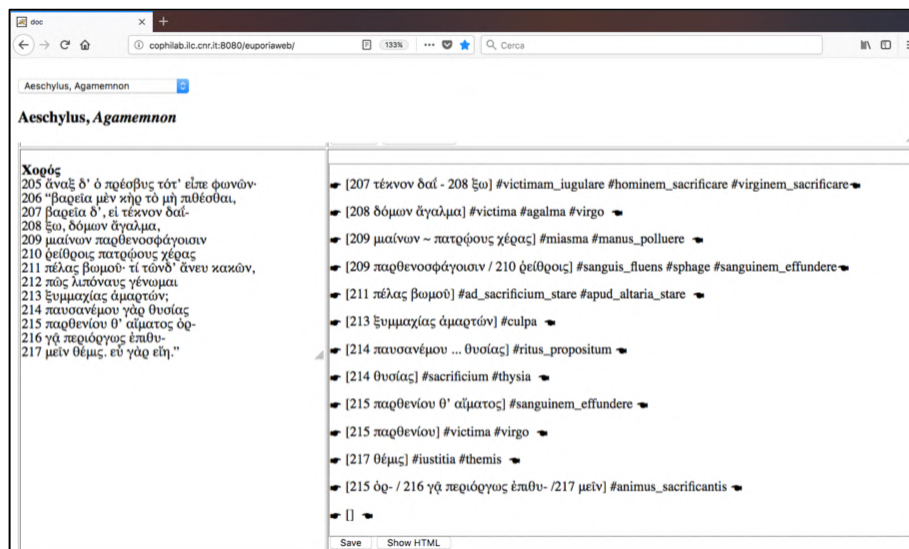


Figura 1: Interfaccia web di EuporiaRAGT<sup>11</sup>

Il processo di annotazione ha seguito un approccio *bottom-up*<sup>12</sup>, senza un set di tag imposto a priori. L'insieme dei tag risultante è stato successivamente organizzato in un'ontologia, che tenesse conto delle dinamiche dell'azione rituale.

L'approccio *bottom-up* ha permesso sia di avere un database di annotazioni molto vicine al dato testuale, sia di strutturare l'informazione seguendo i fondamenti teorici dell'antropologia del mondo antico.

Il *retrieval* delle informazioni è ad oggi possibile tramite *query* SQL sul database delle annotazioni. È inoltre possibile formulare ricerche complesse, attraverso *query* SPARQL sull'ontologia (fig. 2). È prevista la messa appunto di una interfaccia utente che permetta a studiosi di tragedia o religione greca di interrogare il *database* delle annotazioni.

<sup>11</sup> In fase di progettazione, si è scelto di mantenere l'interfaccia più scarna possibile, e di rendere il codice HTML trasparente all'utente, con lo scopo di permettere all'annotatore di restare concentrato soltanto sull'annotazione testuale. La scelta di esprimere gli *hashtag* in latino mira a rendere le annotazioni concise, e avvicina il sistema EuporiaRAGT al progetto *Memorata Poetis*, con il quale si propone di essere compatibile ([www.memoratapoetis.it](http://www.memoratapoetis.it)), vd. Boschetti et al. 2014.

<sup>12</sup> Nell'ambito del progetto *Memorata Poetis*, sono stati annotati temi e motivi negli epigrammi in Greco, Latino, Arabo e Italiano, seguendo un approccio *top down*. Le parole chiave per l'annotazione dei temi e motivi sono state ricavate, in un primo momento, dalla tassonomia latina di un *index rerum notabilium* mentre, in un secondo momento, la tassonomia è stata ampliata e ristrutturata in un'ontologia. Per la ristrutturazione del set dei tag di *Memorata Poetis* nell'ontologia vd. Khan et al. 2016.

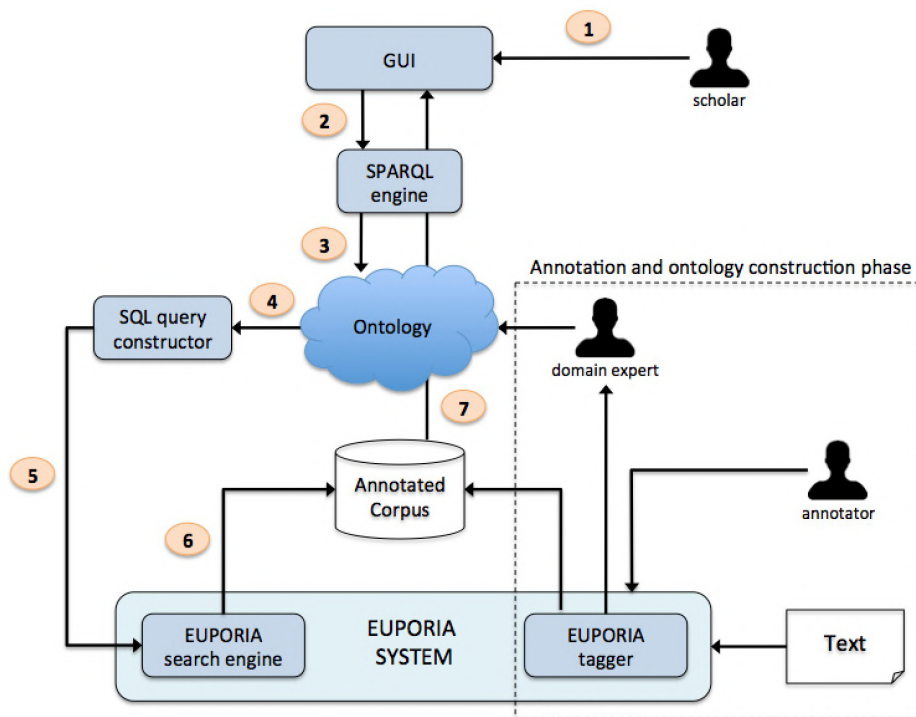


Figura 2: il sistema EuporiaRAGT

## EUPORIA *LanGT* (LANDSCAPES IN GREEK TRAGEDY)

Oltre a un sistema di annotazione centrato sull'utente, Euporia consente di compiere annotazioni collaborative sullo stesso *corpus* di testi: se il risultato dell'annotazione svolta su EuporiaRAGT va a costituire una *personomy*, da un lavoro di annotazione cooperativo si otterrebbe una vera e propria *folksonomy*, che amplierebbe le possibilità offerte dalla strutturazione dei tag nell'ontologia.

Il *corpus* tragico sarà annotato, con principi e obiettivi diversi, anche nell'ambito del progetto Euporia*LanGT*, secondo progetto dottorale nato dalla collaborazione tra il CoPhiLab e il LAMA. L'annotazione, che si concentra sul campo semantico degli spazi naturali, è portata avanti nell'ambito di una ricerca dottorale sul paesaggio naturale nella tragedia attica di V secolo a.C.

Il progetto prevede un'annotazione semantica e tematica del *corpus* tragico, che vada a marcare: gli elementi del paesaggio tragico che hanno rilievo sul piano del rito; le diverse divinità coinvolte nella pratica di culto e che manifestano il loro potere attraverso specifici elementi naturali; le forme del pensiero religioso di età classica in base alle quali un determinato spazio naturale risulta associato a un dio.

In questo caso, il processo di annotazione è semi-strutturato, incentrato su tre tag fondamentali:

- #evidence, marca un elemento o uno spazio naturale esplicitamente menzionato nel testo tragico.
- #ontology, asserisce una relazione ontologica (ad esempio, il rapporto tra una divinità e uno specifico elemento del paesaggio).
- #association, rimanda al piano delle associazioni mentali operate dagli spettatori.

Lo scopo del lavoro è quello di registrare, a partire dalla menzione di un paesaggio naturale (#evidence), le diverse associazioni mentali possibili con sfere di culto afferenti a più divinità (#association), rendendo in questo modo giustizia alla complessità del pensiero religioso dei Greci del V secolo, spettatori della tragedia.

Entrambi i DSL permettono infine di segnalare *variae lectiones* e diverse interpretazioni del testo, con i corrispondenti riferimenti ai manoscritti e/o alla letteratura secondaria.

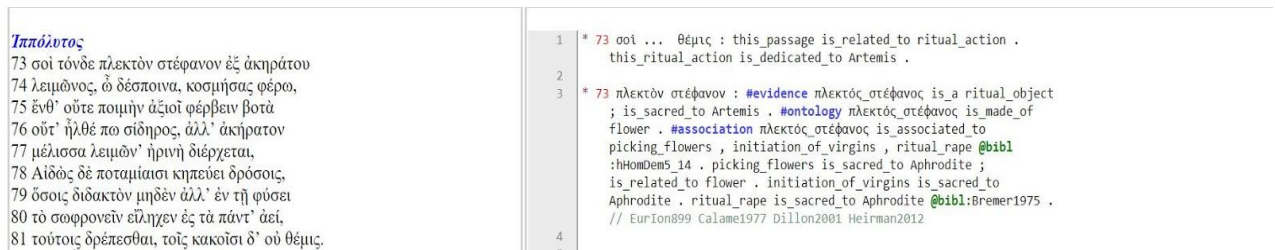


Figura 3: Interfaccia web di Euporia LanGT

## PRIMI ESPERIMENTI DIDATTICI

Oltre a essere adottato per le attività di ricerca del Laboratorio di Antropologia del Mondo Antico, il sistema Euporia è stato utilizzato per alcuni esperimenti di didattica nell’ambito del Laboratorio di Lingua Greca del corso di laurea in Lettere.

Nell’anno accademico 2016/2017, l’annotazione digitale è stata compiuta sul Canto XXIV dell’*Odissea*. Agli studenti è stato chiesto di utilizzare parole chiave funzionali alla traduzione e alla comprensione del testo: sono state annotate particolarità metriche, linguistiche e stilistiche, oltre alla presenza di formule nel testo.

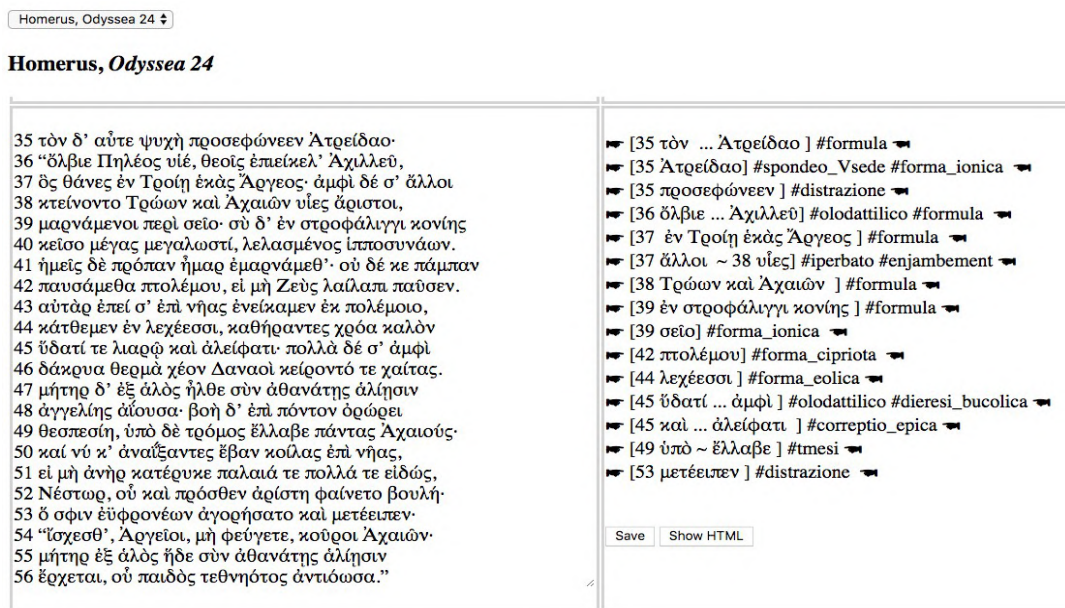


Figura 4: annotazione morfologica e stilistica

Il processo stesso di annotazione si è rivelato particolarmente utile per gli studenti: oltre a comportare un’attività di *close reading* del testo omerico, la selezione delle parole chiave e dei passi da annotare ha permesso agli studenti di concentrarsi su fenomeni da loro ritenuti rilevanti per la corretta comprensione e per la traduzione del passo. L’annotazione va quindi a integrare le informazioni morfologiche e sintattiche ricavabili dalle varie risorse disponibili online (primi tra tutti il database del *Perseus project* e gli strumenti di lettura dei testi antichi offerti dal progetto *Alpheios*) in quanto non costituisce un’analisi sistematica del testo, ma una raccolta di casi notevoli significativi per studenti che hanno appena iniziato un percorso di apprendimento del greco antico a un livello avanzato.

Nell'anno 2017/2018 è stato proposto un lavoro di annotazione che combinasse diversi generi di questioni interpretative, relative alla seconda parte dell'*Aiace* di Sofocle, in cui si mette in scena, intorno al cadavere dell'eroe presente nell'*orchestra*, la discussione sul funerale di Aiace. In particolare, si è proposta agli studenti un'annotazione che registrasse i problemi relativi al cadavere come oggetto scenico e come oggetto rituale, e che tenesse allo stesso tempo conto della funzione sintattica dei riferimenti al corpo di Aiace come agente, e come oggetto delle azioni dei personaggi sulla scena.

I tag creati, anche in questo caso con un approccio completamente *bottom-up*, si sovrappongono in parte all'annotazione rituale e scenica di *EuporiaRAGT*, mentre in parte riguardano la struttura sintattica e le particolarità morfologiche del testo. In alcuni casi, gli studenti hanno aggiunto annotazioni che riguardano lo stato del testo (espunzioni in particolare), riferimenti intertestuali o riferimenti agli scoli.

Soph. Ajax, 990-1016	EuporiaRAGT	EuporiaAjax
<p><b>Χορός</b> καὶ μὴν ἔτι ζῶν, Τεῦκρε, τοῦδ' ἐσσι μέλει ἐφίεθ' ἀνὴρ κείνος, ὥσπερ οὖν μέλει.</p>	<p>☞ [991 ἐφίεθ' ἀνὴρ κείνος] #episkepsis ☞ ☞ [] ☞</p>	<p>☞ [990 καὶ ... 991 κείνος] #voluntas_mortui ☞ ☞ [] ☞</p>
<p><b>Τεῦκρος</b> ὦ τῶν ἀπάντων δὴ θεαμάτων ἐμοὶ ἀγλιστον ὦν προσεῖδον ὀφθαλμοῖς ἐγώ, ὀδὸς θ' ὀδῶν πασῶν ἀνιάσασα δὴ μάλιστα τοῦμόν σπλάγχχνον, ἦν δὴ νῦν ἔβην. ὦ φίλτατ' Αἴας, τὸν σὸν ὡς ἐπησθόμην μόρον διώκων κάξιχνοσκοπούμενος. ὄξεϊα γὰρ σοῦ βάξις ὡς θεοῦ τινοῦ διήλθ' Ἀχαιοὺς πάντας ὡς σίχει θανῶν. ἀγὼ κλύων δύστηνος ἐκποδῶν μὲν ὦν ὑπεστέναζον, νῦν δ' ὀρώων ἀπόλλυμαι. οἴμοι. ἴθ', ἐκκάλυψον, ὡς ἴδω τὸ πᾶν κακόν. ὦ δυσθέατον ὄμμα καὶ τόλμησ πικρᾶς, ὄσας ἀνίας μοι καταστειράς φθίνεις. ποῖ γὰρ μολεῖν μοι δυνατόν, εἰς ποίους βροτούς, τοῖς σσίς ἀρήξαντ' ἐν πόνοισι μηδαμοῦ; ἦ ποῦ με Τελαμών, σὸς πατὴρ ἐμός θ' ἄμα, δέξαιτ' ἂν εὐπρόσωπος ἰλευς τ' ἴσως χωροῦντ' ἄνευ σοῦ. πῶς γὰρ οὐχ; οἶτω πάρα μηδ' εὐτυχοῦντι μηδὲν ἦδιον γελᾶν. οὐτός τί κρύψει; ποῖον οὐχ ἔρεϊ κακόν τὸν ἐκ δορὸς γεγῶτα πολεμίου νόθον, τὸν δειλία προδόντα καὶ κακανδρία σέ, φίλτατ' Αἴας, ἡ δόλοισιν, ὡς τὰ σά κράτη θανόντος καὶ δόμουσ νέμοιμ σούσ.</p>	<p>☞ [ 992 ὦ τῶν ἀπάντων ...1039 στεργέτω κάγω τάδε] #s #apud_corpus_stare ☞ ☞ [ 992 ὦ τῶν ἀπάντων ...993 ἀγλιστον] #s #aspectus ☞ ☞ [995 σπλάγχχνον] #viscera #motus_animi ☞ ☞ [1001 ὑπεστέναζον] #gemitus ☞ ☞ [1003 ἴθ', ἐκκάλυψον] #s #corpus_detegere ☞ ☞ [1003 ὡς ἴδω τὸ πᾶν κακόν] #s #aspectus #mortui_corpus ☞ ☞ [1004 ὦ δυσθέατον ὄμμα] #s #aspectus #mortui_corpus ☞ ☞ [1008 σὸς πατὴρ ἐμός θ' ἄμα] #propinquitatis_vinculum ☞</p>	<p>☞ [992 Τεῦκρος] #s #apud_corpus ☞ ☞ [992 ὦ-δὴ] #gemitus ☞ ☞ [993 ὀφθαλμοῖς] #aspectus ☞ ☞ [992 ὦ...995 ἔβην] #lamentatio #singularis ☞ ☞ [996 ὦ φίλτατ' ...1001 ἀπόλλυμαι] #dolor ☞ ☞ [998 ὡς θεοῦ τινοῦ] @cfr: 998_1 #824_830 ☞ ☞ [1001 ὑπεστέναζον] #dolor #submissus ☞ ☞ [1002 οἴμοι] #gemitus ☞ ☞ [1003 ἴθ', ἐκκάλυψον ... ὡς ἴδω τὸ πᾶν κακόν] #corpus_patiens ☞ ☞ [1003 ἴθ'] {@schol_L} #coryphaeuo ☞ ☞ [1003 ἴθ'] {@schol_L} #Tecmessa_absens ☞ ☞ [1003 ἐκκάλυψον] #s #corpus_manifestus #avato ☞ ☞ [1008 με] @vl:1008_1 Küster ☞ ☞ [1008 ἦ... 1010 σοῦ] @cfr:1008_1 Radt, TGF, 4.576 ☞</p>

Figura 5: annotazione rituale ed esperimento didattico a confronto

Come mostra la fig. 5, l'annotazione dialoga con l'annotazione rituale e scenica di *EuporiaRAGT*: le due annotazioni possono essere confrontate e integrate per una migliore comprensione del testo drammatico.

## VERSO UN'ANNOTAZIONE COLLABORATIVA: I PROGETTI IN CORSO

Nel primo semestre del 2018 è stato avviato un progetto che vedrà impegnati, fino a giugno 2019, sei studenti di Letteratura Greca del primo anno. Si tratta di una annotazione tematica dell'*Agamemnone* di Eschilo, che rientra nella programmazione d'esame.

Ai partecipanti è stato chiesto di scegliere autonomamente una traccia intorno alla quale inserire le proprie annotazioni : è stato così possibile valutare le conoscenze pregresse su alcuni dei temi portanti della tragedia e regolare, così, il livello di complessità del lavoro.

Sono stati selezionati argomenti di carattere morfologico e semantico: forme dialettali doriche particolarità linguistiche e *hapax* nelle sezioni corali; lessico della colpa, della guerra, del sangue e della morte; epiteti *gender-inverted*.

Gli studenti lavorano da soli o in coppia sullo stesso *topic*, senza la possibilità di accedere ai piani di marcatura dei loro colleghi. Non soltanto il lavoro svolto in coppia favorirà l'apprendimento fra pari, ma una volta ultimata l'annotazione sarà possibile confrontare i risultati in un ambiente collaborativo, e unire i diversi *tagset* in una *folksonomy*.

Dal punto di vista didattico, l'esercizio su Euporia sarà utile agli studenti per apprendere più rapidamente il lessico, lo stile e le varianti dialettali del linguaggio eschileo. Allo stesso tempo, il docente potrà valutare i tempi e le modalità di apprendimento, modulando di conseguenza la sua attività didattica.

Dopo alcuni anni di sperimentazione universitaria, sono stati coinvolti nel progetto anche studenti e insegnanti dei licei italiani. In particolare, nel 2018 è stato coinvolto liceo classico di Siracusa, e - a partire dall'a.s. 2018/2019 - saranno coinvolti anche il liceo classico di Pisa e di Venezia. I docenti potranno selezionare i testi da analizzare, e concordare con gli studenti il linguaggio e gli obiettivi dell'annotazione: la creazione di un DSL permetterà di avvicinare la marcatura digitale al lavoro sul testo normalmente svolto in classe. Allo stesso tempo, la classe costituirà un vero e proprio ambiente cooperativo, in cui gli studenti (guidati dal docente) potranno collaborare all'analisi e all'interpretazione del testo, e confrontare i loro risultati.

Il progetto permetterà di osservare, su base statistica ma sull'intero territorio nazionale, la ricaduta locale dell'utilizzo di strumenti digitali e consentirà così di riflettere anche su eventuali *best practices* adottate in diverse aree d'Italia, nell'auspicio di un coinvolgimento sempre più ampio di altri licei.

## BIBLIOGRAFIA

- [1] Alvoni G. 2002. *Scienze dell'antichità per via informatica. Banche dati, Internet e risorse elettroniche nello studio dell'antichità classica*, CLUEB: Bologna.
- [2] Boschetti, F., Del Gratta, R. e M. Lamé. 2014. *Computer Assisted Annotation of Themes and Motifs in Ancient Greek Epigrams: First Steps*, In: Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014, Pisa, 9-10 December 2014, (a cura di) R. Basili, A. Lenci, B. Magnini, Pisa University Press: Pisa, pp. 83-86.
- [3] Condello F. 2018. *La scuola giusta. In difesa del liceo classico*. Mondadori: Milano.
- [4] Fabrini P. 1999. *La rete per gli antichisti*, SEU: Roma.
- [5] Favini L. 2006. *Un modello, un metodo*. In *Nuova Secondaria* n. 1 (settembre 2006), pp. 79-84.
- [6] Gibbs, F., Owens, T. 2012. *Building better digital humanities tools: toward broader audiences and user-centered designs*, In *DHQ* 6, 2 (2012).
- [7] Khan, A. F. et al. 2016a. *Restructuring a Taxonomy of Literary Themes and Motifs for More Efficient Querying*. In *MATLIT* 4 (2), Coimbra 2016a, pp. 11-27.
- [8] Khan, A. F. et al. 2016b. *Leveraging a Narrative Ontology to Query a Literary Text*. In *OASICs-OpenAccess Series in Informatics*, 2016b, pp. 10:1-10:10.
- [9] Mastandrea, P. 2009. *Gli archivi elettronici di Musisque deoque. Ricerca intertestuale e cernita fra varianti antiche (con qualche ripensamento sulla tradizione indiretta dei poeti latini)*. In *Poesia latina. Nuova e-filologia. Opportunità per l'editore e per l'interprete*, (a cura di) P. Mastandrea e L. Zurli. Roma 2009, pp. 41-72.
- [10] Mugelli, G., Boschetti, F. et al. 2016. *A user-centred design to annotate ritual facts in ancient Greek tragedies*. In *BICS* 59 (2), pp. 103-20.
- [11] Mugelli, G., Khan A. F., et al. 2017, *Designing an Ontology for the Study of Ritual in Ancient Greek Tragedy*. In *Proceedings of Language, Ontology, Terminology and Knowledge Structures Workshop (LOTKS 2017)*, Montpellier, pp. 96-105.
- [12] Parr, T. 2010. *Language implementation patterns: create your own domain-specific and general programming language*, Raleigh:NC 2010.

- [13] Roscalla F. 2016. *Greco, che farne? Ripensare il passato per progettare il futuro. Manuali e didattica tra Sette e Novecento*, ETS: Pisa 2016.
- [14] Taddei, A. 2011. *Testi e contesti. Per una didattica della letteratura greca nei nuovi licei*, ETS: Pisa 2011.

# Integrazione di dimensioni narrative e modalità immersive di esplorazione interattiva in visite educative museali

Giovanni Luca Dierna<sup>1</sup>, Alberto Machi<sup>1</sup>

<sup>1</sup> Istituto di Calcolo e Reti ad Alte Prestazioni - CNR, Italia

## ABSTRACT

Il presente lavoro risponde alle esigenze di un potenziamento della didattica museale attraverso gli strumenti della *digital education*. L'articolo descrive in particolare l'attività di sperimentazione di un applicativo su terminale mobile che integra la tecnica dello storytelling con tecnologie di localizzazione indoor nell'ambito di percorsi didattici di conoscenza e contestualizzazione di reperti museali.

Ogni percorso sviluppa un tema culturale, utilizza e integra tra di loro l'attività in aula e la visita museale, con l'ausilio di tecnologie di editoria elettronica su web, visite virtuali, tecnologie di restituzione su terminale mobile, di posizionamento indoor tramite QR-code, ed un'infrastruttura tecnologica composta da radiofari Bluetooth (BLE beacon) posizionati in prossimità di reperti notevoli del museo.

Il percorso didattico tematico inizia a scuola attraverso una visita virtuale del museo lungo il percorso di visita suggerito per familiarizzare con i luoghi, ed una prima introduzione culturale al tema della visita per familiarizzare con la terminologia descrittiva dei reperti. Il percorso prosegue con una visita al museo e la ricerca in modalità ludica di reperti che testimoniano degli usi e costumi raccontati in una storia svelata progressivamente sul terminale mobile. Ogni tappa del percorso prevede una introduzione alla ricerca, una descrizione del reperto, una descrizione del suo contesto d'uso, domande di rinforzo dell'apprendimento a punti. Il percorso culturale si conclude in classe via web con la rivisitazione virtuale degli itinerari e lo studio delle pagine di approfondimento.

Nella sperimentazione in corso tre classi di ragazzi di età fra i 9 ed i 16 anni hanno verificato presso il Museo Archeologico Regionale "A. Salinas" di Palermo la funzionalità della applicazione mobile su due percorsi tematici su usi e costumi nelle colonie greche in Sicilia: "Un giorno da campione" e "Dalla tabula al tablet". Sono emersi suggerimenti sulla usabilità della applicazione, in particolare sul dimensionamento dei testi, sulle modalità di visualizzazione della mappa. La sperimentazione con gli insegnanti del percorso didattico completo alla data di scrittura di questo abstract è in corso.

## PAROLE CHIAVE

Storytelling, caccia al reperto, QR-code, BLE beacon, localizzazione di prossimità.

## INTRODUZIONE

Presenti da alcuni anni in ambito museologico e museografico, anche se ancora in fase sperimentale, e ampiamente praticate e diffuse nel mondo anglosassone, sono le metodologie orientate all'inclusione partecipativa delle collettività dei fruitori e basate sulla *narrazione* o *storytelling*.

E' oramai noto infatti l'uso dello storytelling, ovvero l'uso della tecnica della narrazione, in ambito museale, al fine di costruire un contesto di fruizione immersivo ed emotivo e utilizzato pertanto in ambito didattico e pedagogico per l'apprendimento di bambini e ragazzi. Ed è proprio per tale ragione che, nell'ottica del cosiddetto *audience development*, le iniziative

legate allo *storytelling* sono entrate a far parte dei “servizi aggiuntivi” di musei, insieme al bookshop o al bar/ristorante o alle classiche visite guidate.

In questo senso, le moderne tecnologie assumono un ruolo importante nella fruizione dei beni culturali: gestione di big data e open data, analytics, tecnologie informatiche e multimediali, nuove modalità interattive con i BLE beacon, ricostruzioni virtuali, realtà aumentata, ecc. [1, 2].

Lo *storytelling* si suddivide in “diretto” (il museo si racconta), “indiretto” (i visitatori raccontano il museo) e “partecipativo” (forma mista delle prime due). Quest’ultima si rivela solitamente la più efficace ed è propria questo approccio che qui viene utilizzato in ottica di ricerca, esplorazione e analisi dei reperti museali [3].

Nel presente articolo viene in particolare mostrato un percorso didattico di apprendimento di contenuti legati a un tema di cui i reperti museali sono elementi rappresentativi e con esso viene discussa una sperimentazione legata a visite museali guidate che integrano una dimensione narrativa e una modalità interattiva di esplorazione dei reperti con l’ausilio di tecnologie di localizzazione all’interno di edifici (immagini sferiche con hotspots, radiofari Bluetooth low energy, QR-code).

## IL PERCORSO DIDATTICO

Il percorso didattico qui esposto si inserisce all’interno di strategie messe in atto in ambito museale per favorire una attiva e concreta partecipazione alla produzione e condivisione di significati da parte degli stessi destinatari del messaggio culturale che i musei intendono veicolare.

Esso è stato sviluppato all’interno di un Progetto relativo alla sperimentazione di tecnologie informatiche per la ludicizzazione didattica avviato nel 2017, denominato “Archeogame al Salinas” e finanziato dall’Assessorato dei Beni Culturali e dell’Identità Siciliana della Regione Sicilia.

Il Progetto ha visto il contributo dell’Istituto di Calcolo e Reti ad Alte Prestazioni del Consiglio Nazionale delle Ricerche (ICAR-CNR) - sede di Palermo -, del Museo Archeologico Regionale “Antonino Salinas” di Palermo, museo prescelto per la sperimentazione, e dell’Associazione Culturale “Incontrosenso” di Palermo.

Esso si articola in tre differenti momenti:

- 1) preparazione in classe pre-visita;
- 2) visita museale;
- 3) post-visita.

Nella prima fase, condotta in classe, viene condotto uno studio preliminare con l’ausilio dell’insegnante consistente nello studio degli ambienti del museo e dei reperti ivi esposti. In tale contesto, si introducono le informazioni generali sul museo, si studia il/i percorso/i di visita museale, ecc. In tale contesto può essere condotta anche la visita virtuale del sito [4].

Nella seconda fase viene condotta la caccia al reperto con visita tematica presso il museo [5], con l’ausilio delle tecnologie (applicativo web su tablet, denominato *Archeogames*, QR-code [6] e radiofari Bluetooth) [7] in gruppi di 4-5 ragazzi.

Nell’terza fase, condotta anch’essa in classe, vi è l’elaborazione critica dei contenuti dei percorsi tematici e condivisione delle informazioni, con rivisitazione via web dei percorsi effettuati e lettura degli approfondimenti.

Entrando nel dettaglio, le strategie didattiche utilizzate nella seconda fase sono:

- 1) *approccio narrativo*. Esso si estrinseca nella costruzione di “percorsi narrati” di visita alle collezioni museali (nella fattispecie “Un giorno da campione” sullo sport nell’antichità e “Dalla tabula al tablet” sulla scrittura nel mondo antico). Ciascun percorso è formato da 12 tappe e per ogni tappa è individuato un oggetto rappresentativo che sarà oggetto della ricerca esplorativa;



- 2) *ricerca esplorativa*. Essa si estrinseca in una ricerca sequenziale, vincolata ad un ben preciso itinerario, oppure di libera esplorazione e si avvale di suggerimenti testuali e/o audio e/o video e/o multimediali (cruciverba, puzzle);
- 3) *approccio partecipativo*. Questa strategia, basata su domande a risposta multipla cui occorre rispondere, con relativa acquisizione di punteggio, crea un ponte ed una interazione tra gli utenti e gli oggetti, le tappe e l'intero percorso narrativo;
- 4) *interazione utente-oggetto*. Questa, fondamentale per accertare l'avvenuta e corretta scoperta dell'oggetto cercato, avviene mediante lettura diretta dei QR-code associati all'oggetto o notifica di opportuni messaggi da parte dei radiofari BLE beacon.

## METODOLOGIE ADOTTATE

I contenuti base, le animazioni, i media di corredo e le pagine di approfondimento alle storie sono predisposti cooperativamente dagli autori e dall'editore tramite portale web su un Sistema di gestione di contenuti (CMS). Sono quindi integrati in un ipermedia html5 scaricabile come archivio compresso sulla memoria locale dei tablet e visualizzabile tramite il browser locale del tablet. Una procedura di animazione gestisce l'interazione con l'utente e l'avanzamento del gioco visualizzando ad ogni passo solo i contenuti pertinenti.

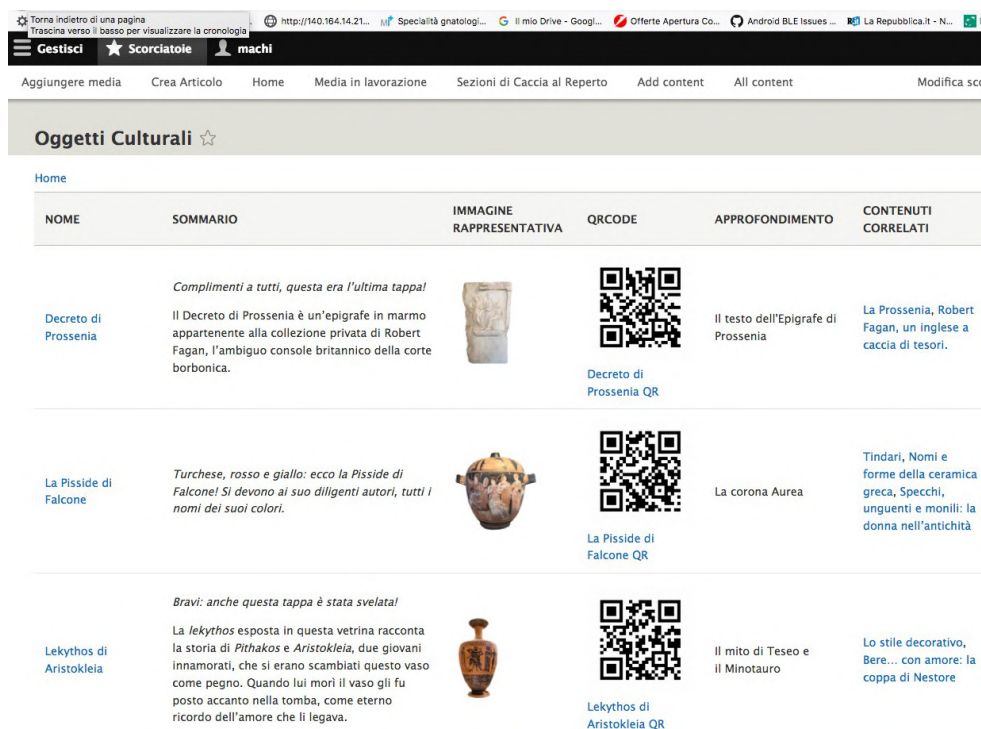


Fig. 1. Schermata di una delle viste riassuntive dei contenuti sul CMS

Ogni storia tematica è usufruibile sia in modalità *itinerario guidato*, con le tappe ordinate e visualizzate in sequenza, o nella modalità *esplorazione* in cui è possibile scoprire le tappe in maniera casuale.

I percorsi tematici sono predisposti partendo da una suddivisione in dodici tappe, ciascuna rappresentata da un oggetto significativo da cercare e scoprire al museo.

La ricerca viene effettuata con l'ausilio di:

1. un indizio audio, video, testuale o anche interattivo-multimediale (cruciverba da risolvere, puzzle da comporre);

2. una mappa su cui vengono campite aree estese di ricerca dei reperti, o indicate le posizioni e lo stato di analisi dei reperti trovati.

La scoperta viene effettuata con l'ausilio di:

1. QR-code, posti in corrispondenza degli oggetti;
2. radiofari Bluetooth low energy, con indicazione sulla mappa di aree di vicinanza sempre più ristrette, in funzione del livello di segnale percepito. I radiofari emettono sempre i loro segnali, ma l'applicazione filtra solo quelli relativi al radiofaro associato alla tappa corrente.

Entrambi permettono di capire se l'oggetto identificato è quello corrispondente all'indizio selezionato. Nel dettaglio, la visita museale avviene attraverso le seguenti fasi:

1. "lettura" dell'indizio (o ascolto dell'audio connesso) e orientamento su un percorso suggerito su una mappa;
2. ricerca dell'oggetto, secondo la metafora del coinvolgimento ludico in una "caccia al tesoro";
3. notifica di avvenuta scoperta tramite lettura del QR-code o messaggio dal beacon, con acquisizione del relativo punteggio;
4. lettura della descrizione con invito all'osservazione del reperto;
5. risposta alla domanda sul reperto trovato, con acquisizione del relativo punteggio;
6. lettura della descrizione del contesto di riferimento del reperto (perché il reperto è associato al soggetto della tappa);
7. risposta alla domanda sulla tappa, con acquisizione del relativo punteggio;
8. lettura dell'approfondimento sul reperto/tappa;
9. lettura dell'intera storia tematica;
10. risposta a tre domande sull'intera storia, con acquisizione del relativo punteggio;
11. "incoronazione" del gruppo vincitore;
12. consegna degli attestati di partecipazione;
13. foto ricordo;
14. saluti finali.

Gli utenti sono assistiti durante la visita tematica dalla app *Archeogames*. La app gestisce il browser, la lettura dei codici QR in prossimità dei reperti e filtra i messaggi dei radiofari, mostrando sulla mappa il livello di prossimità al reperto cercato nella tappa corrente.

Preinstallata e disponibile nei tablet messi a disposizione dal museo, essa permette di accedere a diverse funzioni: dal menu principale si può accedere in particolare alle sezioni "Museo" (con informazioni generiche sul museo), all'introduzione sul percorso tematico scelto, alla mappa del luogo con il percorso suggerito, alle istruzioni e infine alle due modalità di esplorazione (modalità ricerca sequenziale in un itinerario prestabilito e modalità ricerca libera).

Concentrandosi sulla parte dedicata alla vera e propria visita, l'app si configura come un navigatore tra le tappe del percorso individuate tramite un oggetto rappresentativo ad esse collegato.

Il tutto è corredato di indizi per la ricerca, schede informative, domande di verifica comprensione contenuti e sistema di acquisizione punteggi.

Beacon e QR-code posizionati in corrispondenza degli oggetti prescelti hanno il compito di notificare e confermare l'avvenuta scoperta e individuazione dell'oggetto.

Tutte le informazioni sono anche accessibili via web, tramite una versione web dedicata.

L'interfaccia è costruita in maniera tale da consentire la ricerca a tappe obbligate nella modalità "itinerario" e la ricerca random degli oggetti nella modalità "ricerca libera".

La scelta di mettere a disposizione l'applicativo in tablet del museo non pone problemi di archiviazione dei contenuti nei dispositivi mobili dei visitatori, non sempre possibile e agevole a causa della presenza non assicurata della rete wi-fi in tutti i locali del museo.

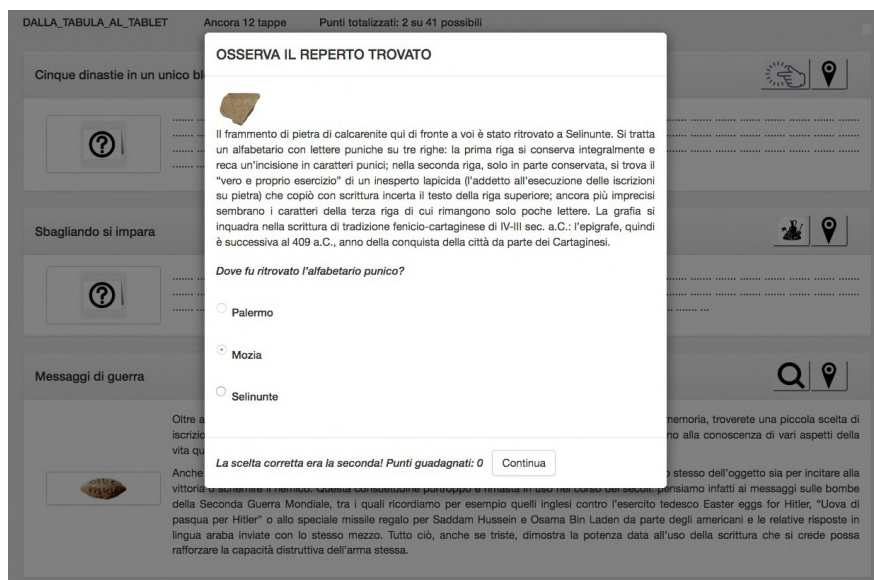


Fig. 2 Schermata della applicazione con la visualizzazione di contenuti relativi a tappe con diverso livello di elaborazione (in modalità esplorazione)

## TECNOLOGIE UTILIZZATE

- La tecnologia di supporto utilizzata nel progetto riguarda:
- progettazione di tecnologie web;
- progettazione di un applicativo per smartphone e tablet con sistema operativo Android;
- progettazione e uso di una piattaforma informatica di editoria elettronica che supporta la realizzazione cooperativa via portale web di cataloghi e guide per Beni ed Eventi Culturali, denominata "Fascicolo Culturale Elettronico", che ha visto il contributo di redattori, revisori ed editori;
- utilizzo di tecniche di lettura diretta (QR-code);
- utilizzo di tecniche di acquisizione indiretta di messaggi provenienti da radiofari Bluetooth;
- approntamento e uso di una infrastruttura tecnologica composta da radiofari BLE;
- uso di terminali mobili (tablet o smartphone) [8].

## DETTAGLI REALIZZATIVI

Il Fascicolo Culturale Elettronico è sviluppato configurando opportunamente un CMS open source Drupal 5.6<sup>13</sup> per la gestione di media e pagine di contenuti multimediali per *Articoli di approfondimento*, *Reperti*, *Sezioni di guide*, *Viste riassuntive* dei contenuti inseriti. Il CMS è dotato di interfaccia di interrogazione per l'estrazione dei dati associati ad ogni singola guida in formato serializzato Json.

Ogni singola guida di itinerario è costituita da un documento ipermediale la cui struttura è descritta in html5, e i cui contenuti sono dinamicamente caricati e visualizzati da uno script

<sup>13</sup> <https://www.drupal.org/>

Javascript che codifica una macchina a stati corrispondenti agli stati di scoperta o approfondimento di ogni tappa del percorso.

Il passaggio fra i vari stati è attivato dalla effettiva pressione di tasti sull'interfaccia da parte dell'utente (sia nell'esecuzione via web che in quella sul tablet) ovvero dalla simulazione della pressione attivata dalla applicazione di controllo al riconoscimento di un codice QR appropriato o della presenza di un beacon nelle vicinanze.

La navigazione virtuale del museo è realizzata utilizzando immagini a 360° in formato *equirectangular* e visualizzate su html5 tramite la libreria javascript open source "*pannellum*"<sup>14</sup>.

I punti di vista significativi per il percorso di visita sono associati ad una area attiva di una mappa del museo. La selezione carica in una finestra grafica un'immagine tridimensionale del punto di vista che contiene a sua volta area attive (hotspots) per la navigazione (tramite ricaricamento di un nuovo punto di vista) o per la attivazione delle informazioni relative al reperto di interesse (vedi fig. 2).

La applicazione di controllo della guida mobile, realizzata in Java su piattaforma mobile Android è costituita da un gestore di 3 attività di interazione con l'utente con interfacce grafiche. L'attività principale consiste nella visualizzazione sulla finestra principale dell'ipermedia guida, utilizzando un browser Chrome, con apertura progressiva delle sezioni pilotata dalla macchina a stati Javascript.

La seconda attività controlla l'acquisizione costante di immagini ed il riconoscimento eventuale di codici QR utilizzando la libreria open source Zxing<sup>15</sup>.

La terza attività interagisce con il sistema operativo Android per attivare la scansione di apparati Bluetooth nelle vicinanze, filtrando quelli dotati dell'identificatore atteso per la tappa corrente.

Le attività secondarie comunicano con l'interprete Javascript del browser Android attraverso una libreria di sistema che permette il richiamo reciproco di procedure fra il browser e le attività Android.

## LA SPERIMENTAZIONE

La sperimentazione effettuata al Museo "Salinas" ha riguardato l'efficacia della metodologia proposta utilizzando sia la versione web che la versione con l'uso di QR-code e beacon.

Essa ha coinvolto tre classi di tre ordini di scuola e ha voluto mettere in evidenza criticità e problematiche nell'uso dell'applicativo. Sono stati così testati:

- 1) il modello di presentazione e rappresentazione;
- 2) l'usabilità dell'applicazione;
- 3) il livello di comprensione dei contenuti;
- 4) la durata della caccia al reperto;
- 5) il grado di partecipazione e di interesse;
- 6) l'efficacia e le ricadute didattiche dell'operazione;
- 7) l'efficacia di utilizzo dei QR-code associati agli oggetti;

Rimangono da verificare:

- 1) la corretta gestione delle soglie di ricezione dei segnali emessi dai beacon;
- 2) l'efficacia della visita virtuale come fase preparatoria alla visita in presenza.

---

<sup>14</sup> <https://pannellum.org>

<sup>15</sup> <https://github.com/zxing/zxing>



**Fig. 3. Uno dei momenti della sperimentazione**

## RISULTATI OTTENUTI

Il test effettuato con le scuole partecipanti ha messo in evidenza<sup>16</sup>:

1. una adeguata modalità di presentazione delle informazioni ,eccezion fatta che per la mappa che è risultata di non immediata intelligibilità da parte degli alunni frequentanti la scuola primaria.

I ragazzi più piccoli (Scuola primaria e Scuola secondaria di I grado) non hanno mostrato abitudine ad utilizzare la mappa per orientarsi, preferendo ricorrere ad una ricerca empirica. La fase di attività preparatoria in classe (visita virtuale) potrebbe rivelarsi un utile esercizio preparatorio;

2. la facilità d'uso dell'applicativo, che si è prestato bene allo scopo;
3. la necessità di rimodulare i testi in funzione del target utente (alunni IV e V classe della Scuola primaria; alunni di Scuola secondaria di I grado; alunni del primo biennio e terzo anno di Scuola secondaria di II grado). La terminologia è invece spesso volutamente tecnica, in accordo al principio di Krashen “input + 1”<sup>4</sup>;
4. la necessità di aumentare il numero di QR-code presenti al museo, al fine di evitare che la ricerca sia focalizzata sulla individuazione di questi ultimi e non sulla ricerca selettiva dei reperti in base alle descrizioni suggerite;
5. la adeguata durata della caccia, compresa fra 45 minuti ed oltre un'ora, in base al numero di approfondimenti visualizzati;
6. un alto grado di coinvolgimento e partecipazione, per tutta la durata dell'operazione;
7. l'efficacia dell'uso dei QR-code.

## CONCLUSIONI

L'utilizzo integrato di tecniche narrative, localizzazione indoor tramite QR-code e beacon rende molto più interattiva e divertente la visita al museo<sup>17</sup>. La possibilità di rivisitare via web i contenuti della visita risulta apprezzata dagli insegnanti, sia in funzione di rinforzo nella acquisizione di nozioni (ad esempio la terminologia tecnica degli oggetti) ma anche per l'integrazione coi percorsi didattici intrapresi nelle scuole.

I musei diventano così non semplicemente luoghi di conservazione delle opere ma sempre più luoghi di apprendimento, in linea con i principi della *gamification*.

<sup>16</sup> <http://www.ildueblog.it/stephen-d-krashen/>

<sup>5</sup> <https://www.youtube.com/watch?v=i7NnIDHjNdw>

## RINGRAZIAMENTI

Gli autori ringraziano:

- l'Associazione Culturale "Incontrosenso" di Palermo, per la redazione dei testi e delle sezioni audio/video e multimediali;
- la dott.ssa Monica Ruffino per il supporto alla sperimentazione didattica;
- il Polo Regionale di Palermo per i parchi e i musei archeologici, Museo Archeologico Regionale "Antonino Salinas" per la supervisione dei testi;
- il Liceo Coreutico Statale "Regina Margherita" di Palermo, l'Istituto Comprensivo Statale "Politeama" – Plesso La Masa – di Palermo e l'Istituto Comprensivo Statale "Guglielmo Marconi" di Palermo, per i test effettuati finora al Museo.

## BIBLIOGRAFIA

- [1] Albano V., Missikoff O., *Nuove tecnologie e beni culturali: domanda e offerta a confronto*, in Granelli A., Tracò F., *Innovazione e cultura. Come le tecnologie digitali potenzieranno la rendita del nostro patrimonio culturale*, il Sole 24 Ore, 2006, Milano.
- [2] Buratti N., 2008 *Nuove tecnologie e percorsi di innovazione in ambito museale*, in "Economia e diritto del terziario", Anno 19 n. 2, Milano, Franco Angeli Editore.
- [3] Calcagno M., Faccipieri S., 2011 *Nuove forme d'interazione nella fruizione di prodotti culturali*, in "Note di Ricerca- Dipartimento di Management", Anno 1 n.1, Venezia, Università Ca' Foscari.
- [4] Bonacasa N., *Il museo online. Nuove prospettive per la museologia*, Oadi, 2011, Palermo.
- [5] Dierna G. L., Machì A. (2017), *Ludicizzazione didattica per la fruizione dei beni culturali. Percorso di "caccia al reperto" museale*. Didamatica17: 31st Didamatica Italian Conference on eLearning - 31a Conferenza Nazionale su Didattica e Informatica, Roma, Italia, 15-16 maggio 2017 (abstract pubblicato sulla rivista "Mondo Digitale", anno XV n. 70, giugno 2017, numero speciale "Tecnologie digitali per rinnovare la didattica l'apprendimento: la sfida per un nuovo modello di sviluppo"), ISBN 9788898091461.
- [6] Besana S. (2010). *L'uso del QR Code come tecnologia didattica: uno studio esplorativo*, in TD-Tecnologie Didattiche n°51, 2010, pp. 34-40.
- [7] Dierna G. L., Machì A. (2017), *Integrazione di QR-code e localizzazione di prossimità in visite educative museali basate su pervasive gaming*. 7th AIUCD Conference 2018. Cultural Heritage in the Digital Age. Memory, Humanities and Technologies, Bari, 31 gennaio – 2 febbraio 2018. Spampinato D. (ed.), AIUCD 2018 - Book of Abstracts, pp. 147-153, ISBN 9788894253528, In: Quaderni di Umanistica Digitale.
- [8] Ceconello M., *Mobile technologies: new ways to access tourism, culture and cities*, Tafter Journal n° 52, 1 ottobre 2012.

# Narrazioni e frammenti, mosaici e algoritmi. La costruzione digitale del racconto tra spazialità e temporalità.

Federico Meschini

Università per Stranieri di Perugia/École Normale Supérieure de Paris

In questa relazione verranno analizzate le forme conoscitive e narrative rese possibili dai social media e in particolare dalla recente funzionalità delle *Instagram Stories*, basate sul concetto di frammento e sull'utilizzo di codici comunicativi eterogenei (Mayer 2001). Il punto di partenza sono le definizioni di tempo e spazio date da Immanuel Kant nella *Critica della ragion pura*, come forme pure della sensibilità, per collegarle alle riflessioni di Lev Manovich sul rapporto tra narrazione e database in cui la temporalità viene ricondotta al concetto di linearità e dei rapporti sintagmatici, e di come nella narrazione questi aspetti siano sempre stati in rapporto dinamico tra di loro, con ricerca continua riconquista della spazialità e di rottura della linearità nel momento in cui quest'ultima si dimostrava un limite, e culminati nella definizione di ipertesto da parte di Theodor Nelson, in cui è proprio l'aspetto spaziale ad essere il principale principio organizzativo (1965). Questo rapporto dinamico si riflette sia a livello di contenuto, come illustrato da Calvino nella lezione sulla molteplicità e conseguente definizione di *iper-romanzo* (1993), sia di forma, con le varie avanguardie e sperimentazioni letterarie in cui la continuità imposta dalla pagina tipografica diventa un limite da superare, come nel caso di *Pale Fire* di Nabokov, in cui la storia si dipana tra il testo del poema e le note, o *Composition nr. 1* di Marc Saporta, il primo *book in a box* in cui ogni singola pagina componente la storia può essere ordinata a piacere dal lettore che riveste perciò un ruolo attivo.

L'aspetto della frammentazione della narrazione insieme alla conseguente necessità di un ruolo maggiormente attivo e consapevole da parte del fruitore sono tra i tratti caratteristici della condizione postmoderna (Lyotard 1979), che secondo Ferraris si è ormai diffusa su larga scala a causa della pervasività dell'informazione digitale, e di conseguenza rende necessarie quelle capacità di decodifica della complessità prima relegate ad un ambito letterario/estetico (Ferraris 2017).

A questo riguardo, riprendendo come esempio la libera disposizione delle pagine nell'opera di Saporta, è facile collegare questa combinazione fattoriale al concetto di complessità computazionale degli algoritmi, in cui ritornano gli aspetti relativi a tempo e spazio, in quanto utilizzati per definire le varie classi di complessità, e al concetto di entropia informativa (Shannon 1948), ossia l'imprevedibilità di un sistema nel generare informazione, tanto maggiore quanto più le varie possibilità delle singole scelte sono indifferenziate tra di loro.

Utilizzando questi concetti, e le relative nozioni necessarie di calcolo combinatorio, ossia i concetti di disposizione semplice e disposizione con ripetizione, partendo dall'*Edge Rank* di Facebook, la prima formalizzazione nella selezione e ordinamento dei singoli elementi costituenti una timeline, verrà analizzato il ruolo degli algoritmi nel compito di riduzione dell'entropia in base a procedimenti probabilistici di *machine learning*, e di come siano le azioni stesse degli utenti a creare il racconto che osservano e di cui in maniera coalescente fanno parte, in quella modalità attiva che non può non ricordare il principio d'indeterminazione di Heisenberg. La fruizione dei contenuti della timeline è stata paragonata da Manovich come un ritorno all'esperienza del *flâneur* (2012), teorizzata da Baudelaire e resa celebre da Benjamin che ne *Les passages parisiens* riesce a descrivere la complessità della metropoli per eccellenza del XIX secolo solo tramite descrizione frammentarie. L'apparente linearità della timeline altro non è che una delle tante possibilità rese possibili dalla logica del

database (Manovich 1999), in quanto permette di esplicitare la forma nascosta di ogni racconto, quelle strutture sottostanti studiate a partire da Propp e dai formalisti russi in numerosi settori culturali, dall'antropologia (Campbell 1949) alla scrittura cinematografica (Vogler 1992).

Infine il paradigma delle storie di Instagram, l'attuale esempio per eccellenza di narrazione tramite frammenti, verrà analizzato utilizzando i modelli sviluppati da Scott McCloud nell'analizzare il medium fumetto (1993), non a caso citato sia da Calvino sia da Manovich in quanto capace di coniugare sia l'aspetto spaziale sia temporale come nessun altro mezzo espressivo, come dimostrato dall'opera di Gianni De Luca, in cui la mancanza della divisione in vignette, ripresa in tempi più recenti da autori come Frank Miller, conferisce alla narrazione una dinamicità altrimenti impossibile da ottenere.

In particolare i concetti utilizzati sono: la *closure*, il ruolo attivo del fruitore nel collegare sinapticamente due informazioni apparentemente disgiunte; le diverse transizioni possibili tra le suddette informazioni, adattate a quello che sono i singoli oggetti discreti disponibili; i possibili rapporti esistenti tra informazioni eterogenee partendo da quello proposto da McCloud tra parole e immagini. Su queste basi verranno mostrati diversi esempi di narrazioni basate sulle *Instagram Stories*, differenziate tra un uso personale e dionisiaco da un lato e relativo alle organizzazioni e apollineo dall'altro, con le diverse possibili interazioni tra questi due principi, soffermandosi infine sulle biblioteche e sul loro utilizzo delle *stories*, con finalità che spaziano dalla semplice promozione delle attività fino al *reference*.

## Bibliografia

- [1] Calvino, I. (1993), *Lezioni americane*. Milano: Mondadori (1 ed. 1988, Milano: Garzanti).
- [2] Campbell, J. (1949), *The Hero with a Thousand Faces*, New York: Pantheon Books.
- [3] Ferraris, M. (2017), *Postverità e altri enigmi*, Bologna: Il Mulino.
- [4] Lyotard J. F. (1979), *La Condition postmoderne. Rapport sur le savoir*, Paris: Éditions de Minuit.
- [5] Mayer, R. (2001), *Multimedia Learning*, New York: Cambridge University Press.
- [6] Manovich, L. (1999), *Database as a Symbolic Form*, in "Millennium Film Journal", nr. 34, <[http://www.mfj-online.org/journalPages/MFJ34/Manovich\\_Database\\_FrameSet.html](http://www.mfj-online.org/journalPages/MFJ34/Manovich_Database_FrameSet.html)>.
- [7] Manovich, L. (2012), *Data stream, database, timeline (part 1)*, <<http://lab.softwarestudies.com/2012/10/data-stream-database-timeline-new.html>>.
- [8] McCloud, S. *Understanding Comics* (1993), Tundra Publishing, Northampton (MA).
- [9] Nelson, T. (1965), *Complex information processing: a file structure for the complex, the changing and the indeterminate*, in *ACM '65 Proceedings of the 1965 20th national conference*, pp. 84-100 <<https://archive.org/details/nelson-file-structure/page/n3>>.
- [10] Vogler, C. (1987), *The Writer's Journey: Mythic Structure for Writers*, Los Angeles: Michael Wiese Productions.
- [11] Shannon C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379–423.



# Il Trattato di Scienza Universal di Vivaldo Belcalzer: fra tradizione e segmentazione testuale. Prove di dialogo ecdotico

Rosa Casapullo, Luciano Longo  
[casapullo@gmail.com](mailto:casapullo@gmail.com); [longo@olomedia.it](mailto:longo@olomedia.it)

## ABSTRACT

Il presente contributo si pone l'obiettivo far dialogare due diverse metodologie ecdotiche, quella tradizionale e quella digitale, su questioni che riguardano la *constitutio textus* del *Trattato di scienza universal* compendio in volgare mantovano del *De proprietatibus rerum* di Bartolomeo Anglico compilato dal notaio Vivaldo Belcalzer. La proposta avrà una doppia articolazione; la prima a cura di Rosa Casapullo sulla storia della tradizione volgare e latina del *Trattato*; la seconda a cura di Luciano Longo sulla costituzione di un modello di marcatura del *Trattato* che cerchi di dar conto sia della tradizione del *Trattato*, sia della tradizione del *De proprietatibus rerum*, ma soprattutto delle relazioni di quest'ultimo con la tradizione volgare. Le due articolazione avranno lo scopo di porre in evidenza problemi ecdotici comuni per i due ambiti di ricerca e tenderanno di presentare in una direzione di sintesi metodologica, procedimenti standardizzati, ad esempio, di formalizzazione del testo base o di rappresentazione delle operazioni di segmentazione testuale. La proposta vuole essere un primo momento di confronto fra due metodologie in una direzione di sintesi con la finalità di produrre una edizione scientifica digitale che sommi in sé i risultati migliori della tradizione ecdotica con quelli più misurati e produttivi dell'ecdotica digitale.

## PAROLE CHIAVE

Testo; copia; metodo; codifica; *mouvance*.

## PROPOSTA

Il presente contributo ha l'obiettivo di presentare alcune questioni metodologiche che riguardano la *constitutio textus* del *Trattato di scienza universal* (d'ora in poi *Trattato*) compendio in volgare mantovano compilato dal notaio Vivaldo Belcalzer del *De proprietatibus rerum* di Bartolomeo Anglico. Le prospettive di indagine saranno due; la prima farà riferimento alla prassi dell'ecdotica tradizionale, la seconda a quella dell'ecdotica digitale. Il *Trattato* è sostanzialmente inedito, se si escludono le pagine edite e commentate da Ghino Ghinassi nel 1965, gli studi avviati nel 2003 da Rosa Casapullo e l'edizione critica nel 2010 dei primi quattro libri a cura della stessa studiosa.<sup>18</sup> La presente proposta avrà una doppia articolazione; la prima a cura di Rosa Casapullo sulla storia del testo o dei testi che sostanziano la tradizione volgare e latina del *Trattato* e che informano l'edizione critica già pubblicata e quella in corso di realizzazione; la seconda a cura di Luciano Longo che, sulla scorta degli studi già avviati, proverà a costituire un modello di marcatura del *Trattato* che cerchi -su passi esemplificativi dei libri XVI, XVII e XVIII- di dar conto sia della tradizione del *Trattato*, sia della tradizione del *De proprietatibus rerum* ma soprattutto delle relazioni di quest'ultimo con la tradizione volgare.

---

<sup>18</sup> È in corso di realizzazione l'edizione critica dei libri XVI, XVII e XVIII del *Trattato di scienza universal* a cura di Rosa Casapullo e Luciano Longo.

Il *Trattato* è testimoniato da quattro manoscritti; il primo, il ms. Additional 8785 della British Library (= L), conserva uno dei più antichi compendi in volgare del *De proprietatibus rerum* redatto dal Belcalzer a fine Duecento. Gli altri tre codici che tramandano il testo conservato in L sono tre copie descritte: il quattrocentesco Riccardiano 2155 (= F), il Canoniciano italiano 24 (=O2) e il Canoniciano italiano 131 (=O1). Dei tre descritti, F e O2 tramandano il testo in modo tendenzialmente fedele, O1 è stato esemplato su O2. In linea di massima, le innovazioni rientrano nella normale casistica degli errori di copia, spesso poligenetici, indotti, soprattutto in F, esemplato da un fiorentino di stanza a Venezia. Nel complesso né F, né O2 presentano un testo modificato intenzionalmente, ma qualche volta lezioni evidentemente guaste dell'originale sono state ritoccate mediante un discreto restauro conservativo. Il ms. più attivo nel rimaneggiare autonomamente il testo è O1, che si colloca al piano più basso della tradizione ed è anche l'unico esemplare sicuramente datato (1466). Solo in O1, infatti, la sintassi è alterata intenzionalmente, il testo è non di rado riassunto e l'ordine delle parole è assai spesso modificato. Il paratesto, inoltre, è completamente riorganizzato: della divisione originaria di L resta poco, mentre sono introdotte *ex novo* varie rubriche. O1 è anche il ms. che con maggior frequenza, rispetto agli altri due, altera erroneamente e banalizza le lezioni tradite, soprattutto il lessico colto. Insomma, se gli errori di F e O2 sono dovuti principalmente a interferenza linguistica, quelli di O1 dipendono in massima parte da un netto dislivello culturale rispetto alla fonte. L'aspetto più rilevante della tradizione del *Trattato* risiede nel fatto che, ancora nella prima metà del Quattrocento il testo è stato recepito e compreso tutto sommato integralmente, con solo un lieve abbassamento di livello, rispetto all'originale, come attestano gli errori relativi alle parti più astratte e ai linguaggi settoriali. Non diversamente dalle enciclopedie in latino, dunque, il *Trattato* è il prodotto di una stretta correlazione fra il testo e la forma che esso assume. Infatti, si può affermare che nel compilare la propria enciclopedia il Belcalzer si tenne stretto a un modello preciso, anche formale, che modificò e adattò a scopi e destinatari diversi. Il *Trattato*, inoltre, ha ereditato dal *De proprietatibus rerum* una duplice scansione: per materia, fino al libro XIII; alfabetica, nei libri XIV-XVIII; mista nel XIX. Tutto sommato l'enciclopedia in volgare mantovano sembra presentare la forma base, la più scarna ed essenziale, fra quelle osservabili nella variegata tradizione del *De proprietatibus rerum*. Come altre enciclopedie mediolatine le quali, sottratte al circuito entro il quale e per il quale erano state originariamente concepite, furono soggette a mutamenti nella struttura, oltre che nella lingua, anche il *Trattato* è stato rielaborato nei contenuti grazie a una drastica selezione e una nuova gerarchizzazione degli argomenti. È necessario aggiungere che il notaio mantovano non è sempre il solo responsabile della semplificazione dei contenuti, della loro banalizzazione o dei cambiamenti sintattici che intervengono nel corso della traduzione. Come si è già accennato, fra la versione standard del *De proprietatibus rerum* e il *Trattato* deve essere ipotizzato un manoscritto (e forse più d'uno) che tramandava un testo già in parte semplificato. Nel *Trattato* la struttura gerarchica del latino è come azzerata; oltre a ciò sono assenti le parti non strettamente funzionali alla linea principale del discorso. La necessità di concentrare il testo latino ha imposto di sciogliere ed eventualmente di appianare le allusioni che si fondano su conoscenze che invece l'enciclopedista latino poteva dare per acquisite, riducendo al minimo la necessità per il lettore di operare inferenze, secondo un modulo tipicamente didattico ben documentato nella tradizione delle enciclopedie in volgare<sup>7</sup>. Se confrontiamo il numero delle informazioni del *De proprietatibus rerum* con quelle del *Trattato*, ci accorgiamo che in quest'ultimo esse sono di numero notevolmente inferiore; la percentuale della riduzione diminuisce notevolmente, però, se confrontiamo le informazioni superstiti nel *Trattato* con quelle delle versioni latine rimaneggiate, variamente presenti nei primi quattro libri del testo in volgare mantovano. È noto che nel *Trattato* mancano cospicue porzioni dell'enciclopedia di Bartolomeo Anglico. Partendo da queste acquisizioni è possibile aggiungere qualche particolare, reso possibile

dalle ultime ricognizioni sul *De proprietatibus rerum* (d'ora in poi: DPR), al fine di chiarire le relazioni fra il *Trattato* e il ramo della tradizione latina, se non il ms., da cui il *Trattato* dipende. Si sa, oggi, che la tradizione del DPR è assai più movimentata di quanto le ricerche aurorali sulla sua circolazione manoscritta lasciassero ipotizzare. Esistono, infatti, redazioni meno diffuse, nelle quali il numero dei libri che compongono l'enciclopedia è maggiore o minore; inoltre sono noti veri e propri rifacimenti, nei quali il testo del DPR risulta completamente rimaneggiato. Il *De proprietatibus rerum* è tramandato, nella tradizione «standard» in diciannove libri (=DPR19), portata a termine verosimilmente entro gli anni Quaranta del Duecento, da circa 200 mss., la gran parte dei quali confezionata entro il secolo XIV. Questa prima fase si accompagna a una vivace attività di rimaneggiamento, il cui apogeo può essere datato all'incirca attorno al 1300. La seconda fase, che coincide con l'attività dei traduttori, comincia poco oltre la metà del Duecento e perdura fino al sec. XV. Se precedentemente tanto i mss. latini quanto i rifacimenti furono prodotti principalmente presso *scriptoria* monastici, in questa seconda fase prevalse l'iniziativa dei poteri regi e signorili laici. L'ultima fase, infine, quella della stampa, si sovrappone all'attività dei traduttori: comincia col 1472, quando fu pubblicata la prima delle oltre 50 edizioni del DPR, e si conclude, nella sua fase antica, col 1609. Fra le redazioni che si discostano dal testo di DPR19 una, in particolare, è utile al confronto col *Trattato*, quella tramandata dai mss. L105 e L147. Questi due testimoni presentano un'evidente somiglianza reciproca nella bipartizione del libro XVI; il testo che tramandano comprende non diciannove libri, ma venti (= DPR20). L105 e L147 condividono integrazioni, lezioni caratteristiche ed errori estranei a DPR19. L105, inoltre, è stato corretto in margine da una mano, diversa da quella del copista, che aggiunge, corregge o integra sulla base del testo standard in diciannove libri. Dunque, la tradizione testuale del *Trattato* si presenta complessa sia sul versante del volgarizzamento sia sul versante del testo latino di dipendenza o derivazione.<sup>19</sup>

La presente proposta, come si accennava prima, avrà una doppia articolazione; la prima a cura di Rosa Casapullo; la seconda a cura di Luciano Longo. La relazione proverà a far dialogare due prospettive, in una direzione di sintesi metodologica, facendole confrontare sui procedimenti di formalizzazione del testo base; sulla risoluzione di situazioni di complessità documentaria; sulla rappresentazione delle operazioni di segmentazione testuale dei manoscritti rispetto al testo di dipendenza latina; sulla varietà delle situazioni testuali dovute a presenza di varianti, errori o addirittura a sezioni escluse nel processo di copia. La ricerca metodologica di confronto si baserà, oltre che sui temi già esposti prima, anche sulla natura instabile della trasmissione di un testo trascritto in copia che genera una serie di errori o di varianti di varia tipologia, ora da antigrafo ad apografo, ora per collazione di una copia con l'altra. L'ecdotica tradizionale di fronte la fluidità testuale di un manoscritto impone all'atto della *constitutio textus* la formalizzazione di un testo archetipico. Invece, la possibilità di sintetizzare alcuni percorsi filologici dell'ecdotica tradizionale con quelli dell'ecdotica digitale possono portare far riflettere più che sulla formalizzazione di un testo archetipico sulla standardizzazione di un testo "vero" per determinati campi referenziali grazie alla naturale intertestualità allargata di un manoscritto prodotto in copia in età medievale-umanistica.

Anche in questa prospettiva nasce la proposta di intervento sull'edizione del *Trattato*. Rispetto a una edizione che faccia emergere giustamente il lavoro di ricostruzione critica del testo tradito stabilito dall'editore, l'esperimento di edizione digitale vuole fondarsi sulla presentazione delle *variae lectiones* significative che possono contribuire ad identificare *in primis* il testo in copia come entità storica e solo dopo come entità intellettuale cioè come prodotto del lavoro ecdotico. Esemplificando il lavoro di confronto, una delle prime questioni

---

<sup>19</sup> Cfr. Vivaldo Belcalzer, *Trattato di scienza universal*, Vol. I: libri I-IV, (a cura di Rosa Casapullo), (Alessandria, Edizioni dell'Orso, 2010), pp. VII-CXX.

che sarà affrontata sarà la scelta tra la definizione di un testo base o la presentazione diacronica e sincronica dei prodotti scrittori storicamente esistenti. Naturalmente questa operazione determinerà una diversa soluzione di marcatura e un diverso modello teorico con un corredo specifico di tag; infatti, se la soluzione adottata sarà quella di considerare il testo-fonte come base per l'edizione critica digitale, la soluzione potrebbe risiedere nella trascrizione-codifica del testo-fonte in un file XML e collegare ad esso l'elenco di tutte le varianti e annotazioni della tradizione volgare. Diversamente, se la soluzione adottata sarà quella di presentare i testi storicamente esistenti e la loro *mouvance* testuale, la soluzione potrebbe risiedere nella trascrizione-codifica di un file XML strutturato con il modulo *Apparatus* della TEI. Una altra questione riguarderà i livelli di complessità delle informazioni metatestuali e paratestuali sia della tradizione volgare sia di questa in dipendenza di quella latina. Una terza questione riguarderà la cosiddetta versione diplomatica dei testi e i livelli di formalizzazione e codifica delle loro testualità. Ultima, ma non meno importante, sarà la tipologia di soluzione da adottare sia a livello di codifica che a livello di rappresentazione per riprodurre i rapporti tra tradizione latina e tradizione volgare sulle azioni di segmentazione testuale.

In conclusione, la proposta vuole essere un primo momento di confronto fra due metodologie in una direzione di sintesi con la finalità di produrre una edizione scientifica digitale che sommi in sé i risultati migliori della tradizione ecdotica con quelli più misurati e produttivi dell'ecdotica digitale.

## BIBLIOGRAFIA

- [1] Altieri Biagi, M.L., 1998. Forme della comunicazione scientifica, in Letteratura italiana, in Ead., Fra lingua scientifica e lingua letteraria, Pisa-Roma-Venezia-Vienna, pp. 21-73.
- [2] Casapullo, R., 2001. "Segmentazione del testo e modalità d'uso delle enciclopedie tra latino e volgare", in *Le parole della scienza. Scritture tecniche e scientifiche in volgare (secoli XIII-XV)*, a c. di Gualdo, R., Galatina, 2001, pp. 153-181.
- [3] Casapullo, R., 2004. Il sistema dei connettivi in alcuni libri del volgarizzamento mantovano del 'De proprietatibus rerum' di Bartolomeo Anglico, in SintAnt. La sintassi dell'italiano antico, Atti del Convegno internazionale di studi (Università di "Roma Tre", 18-21 settembre 2002), a cura di M. Dardano e G. Frenguelli, Roma, pp. 79-100.
- [4] Casapullo, R., 2006. "Le malattie della mente nel volgarizzamento mantovano del De proprietatibus rerum di Bartolomeo Anglico (Libri I-IV, V, VII)", in *Quaderns d'Italia*, 11, pp. 29-53.
- [5] Casapullo, R., 2006. "Il volgarizzamento mantovano del De proprietatibus rerum: il londinese add. 8785 e i suoi descritti quattrocenteschi", in *Lo scaffale della biblioteca scientifica in volgare (secoli XIII-XVI). Atti del convegno, Matera (14-15 ottobre 2004)*, a c. di Librandi R., Piro, R., 2006, pp. 123-140.
- [6] Casapullo, R., Policardo, M.R., 2003. "Tecniche della divulgazione scientifica nel volgarizzamento mantovano del «De proprietatibus rerum» di Bartolomeo Anglico", in *Lingua e Stile*, XXXVIII, pp. 139-176.
- [7] Dardano M., 1967. "Note sul bestiario toscano", in *Dardano* 1992, pp. 37-128.
- [8] Ghinassi, G., 1965. "Nuovi studi sul dialetto mantovano di Vivaldo Belcalzer", in *Studi di filologia italiana*, 23, pp. 19-172.
- [9] Ghinassi, G., 1966. *Il volgare mantovano nell'epoca di Dante*, in *Dante e la cultura veneta, Atti del Convegno di studi (Venezia, Padova, Verona, 30 marzo-5 aprile 1966)*, a c. di Branca, V., e Padoan, G., Firenze, pp. 87-93, pp. 129-135.
- [10] Ghinassi, G., 1976. *Il volgare mantovano tra il Medioevo e il Rinascimento*, in *Ludovico Ariosto: lingua, stile e tradizione, Atti del Congresso organizzato dai comuni di Reggio Emilia e Ferrara, 12-16 ottobre 1974*, a c. di Segre, C., Milano, pp. 7-28.

- [11] Ghinassi, G., 2006. *Dal Belcalzer al Castiglione. Studi sull'antico volgare di Mantova e sul «Cortegiano»*, a c. di Bongrani, P., Firenze.
- [12] Meyer, H., 1988. "Bartholomaeus Anglicus De proprietatibus rerum. Selbstverständnis und Rezeption", in *Zeitschrift für deutsches Altertum und deutsche Literatur*, 117, pp. 237-274.
- [13] Meyer, H., 1991. "Ordo rerum und Registerhilfen in mittelalterlichen Enzyklopädiehandschriften", in *Frühmittelalterliche Studien*, 25, pp. 315-339.
- [14] Meyer, H., 1996. "Die illustrierten lateinischen Handschriften im Rahmen", in *Frühmittelalterliche Studien*, 30, pp. 368-395.
- [15] Meyer, H., 2000. *Die Enzyklopädie des Bartholomäus Anglicus. Untersuchungen zur Überlieferungs- und Rezeptionsgeschichte von "De proprietatibus rerum"*, München, W. Fink.
- [16] Salvioni, C., 1902. "Di un documento dell'antico volgare mantovano", in *Rendiconti dell'Istituto Lombardo di Scienze e Lettere*, s. II, XXXV, pp. 957-970.
- [17] Sánchez González de Herrero, 2007. *De las partes de la tierra y de diversas provincias. Las versiones castellanas del libro XV de "De proprietatibus rerum"*, estudio y edición de M. de las Nieves Sánchez González de Herrero, Vigo, Editorial-Academia de Hispanismo.
- [18] Seymour, M.C., 1975. *On the Properties of Things. John Trevisa's Translation of Bartholomaeus Anglicus De Proprietatibus Rerum. A Critical Text*, voll. I-III, Oxford.
- [19] Seymour, M.C., 1992. *Bartholomaeus Anglicus and his Encyclopedia*, Aldershot-Brookfield.
- [20] Dezza, E., Lorenzoni, A.M., Vaini, M., 2002. "Statuti bonacolsiani", in *Fonti per la storia di Mantova e del suo territorio*, vol. 7 (con un saggio inedito di Torelli P.), Mantova.
- [21] Van den Abeele, B., Meyer, H., 1996. "Die illustrierten lateinischen Handschriften im Rahmen der Gesamtüberlieferung der Enzyklopädie des Bartolomäus Anglicus", in *Frühmittelalterliche Studien*, 30, pp. 368-395.
- [22] Van den Abeele, B., 2005. *Bartholomaeus Anglicus, De proprietatibus rerum. Texte latin et réception vernaculaire – Lateinischer Text und volkssprachige Rezeption*, Brepols.
- [23] Van den Abeele, B., Meyer, H., Ribémont, B., 1999. "Éditer l'encyclopédie de Barthélemy l'Anglais: Vers une édition bi-lingue du De proprietatibus rerum", in *Cahiers de Recherches Médiévales (XIIIe-XVe s.)*, 6, pp. 7-18.

# Diogene alla ricerca dell'uomo contemporaneo: le Digital Humanities "lucerna" per riconoscere il Digital Cultural Heritage. La riflessione DiCultHer-AIUCD

Nicola Barbuti, Fabio Ciotti, Giuliano De Felice

nicola.barbuti@uniba.it fabio.ciotti@uniroma2.it giuliano.defelice@unifg.it

## ABSTRACT

In questo paper si presentano gli esiti di alcune riflessioni maturate negli ultimi cinque anni dalla *Scuola a Rete per la Formazione nel Digital Cultural Heritage, Arts and Humanities – DiCultHer* congiuntamente con l'AIUCD, sul tema del riconoscimento del ruolo che le *Digital Humanities* (DH) hanno nella costruzione di una nuova Cultura Digitale, oggi indispensabile per identificare e certificare il *Digital Cultural Heritage* richiamato nelle *Considerazioni UE* del 21 maggio 2014 sul Patrimonio Culturale. Si rende perciò irrinunciabile una collocazione epistemologica-disciplinare delle DH nel sistema formativo e didattico nazionale ed internazionale quale settore scientifico disciplinare autonomo, per elaborare metodi, tecniche e modelli in grado di contribuire a creare la *Catena del Valore Formativo*, nella quale Scuola, formazione professionale (Istituti Tecnici Superiori) e Università devono concorrere per provvedere le nuove generazioni delle conoscenze e competenze culturali digitali necessarie a riconoscere nelle entità digitali il valore di memoria e di fonte conoscenza che le qualificano come Patrimonio Culturale.

## PAROLE CHIAVE

Cultura Digitale, Digital Cultural Heritage, Digital Humanities, Didattica, Formazione, Processo, Memoria.

## INTRODUZIONE

L'art. 2 delle *Conclusioni del Consiglio dell'UE* del 21 maggio 2014 sul "*Patrimonio culturale come risorsa strategica per un'Europa sostenibile (2014/C 183/08)*" ha inquadrato l'evoluzione che il digitale ha avuto negli ultimi dieci anni, ridefinendo il Patrimonio Culturale come segue:

"Cultural heritage consists of the resources inherited from the past in all forms and aspects – tangible, intangible and digital (born digital and digitized), including monuments, sites, landscapes, skills, practices, knowledge and expressions of human creativity, as well as collections conserved and managed by public and private bodies such as museums, libraries and archives. It originates from the interaction between people and places through time and it is constantly evolving".

In questa definizione ci pare di poter rilevare la consapevolezza che la *rivoluzione digitale* è una fase ormai in conclusione in quanto sta evolvendo in *Era Digitale*, e che questa Era si fonda sul riconoscimento di cosa e quanto possa essere riconosciuto come nuovo *Digital Cultural Heritage* (DCH).

Partendo da questo assunto, diviene urgente ripensare i processi di digitalizzazione e di co-creazione del digitale quali espressioni sociali e culturali dell'epoca contemporanea, nella quale, tuttavia, ancora manca un'autentica *Cultura Digitale* che consenta di individuare approcci e metodi per identificare il *Digital Cultural Heritage* (DCH) tra le entità digitali

create fino a oggi e in produzione, basandosi su criteri chiari e omogenei per validarle e certificarle come memoria e fonti di conoscenza.

Situazione paradossale, ove si pensi che, fin dal principio, alcuni settori disciplinari umanistici hanno mostrato grande interesse nei confronti dell'innovazione digitale. In particolare, le conoscenze *biblioteconomiche* e *archivistiche*, *linguistico letterarie* e *archeologiche* si sono subito confrontate con i nuovi scenari della rivoluzione digitale. Sono state sviluppate esperienze di ricerca e didattica pionieristiche, ma basate per lo più sull'uso strumentale delle nuove istanze del *data science* e del digitale nei campi della gestione della conoscenza, dell'acquisizione dei dati, del supporto all'interpretazione, arrivando a generare linee formative nuove nei settori disciplinari tradizionali, quali *scienze biblioteconomiche e dell'informazione*, *archivistica informatica*, *linguistica computazionale*, *archeologia digitale*.

Oggi però, superata la luna di miele in cui alle tecnologie veniva chiesto (spesso acriticamente) un supporto al miglioramento delle capacità di analisi e di sintesi, le discipline umanistiche sono orientate verso i nuovi temi aperti dalle prospettive delle *Conclusioni UE*, in una dimensione pubblica, condivisa, partecipata, in cui il ruolo del digitale, ancora relegato alla fenomenologia *social*, è ancora tutto da costruire.

## **UN FARO FUORI TEMPO**

Principale riferimento delle nuove prospettive aperte dalle *Conclusioni UE* n. 2 è la *Convenzione di Faro*, documento cruciale che ha innescato, anche in Italia, una profonda rivisitazione del concetto di Patrimonio Culturale, non più da intendere come un insieme di oggetti e beni (che siano essi materiali e immateriali poco importa), ma come una rete articolata di relazioni fra comunità e patrimonio culturale.

Lo scenario risulta particolarmente delicato, visto il ruolo assolutamente marginale riservato al digitale nella *Convenzione*: anacronisticamente considerato in funzione meramente strumentale di valorizzazione del patrimonio tangibile e intangibile, guardato con circospezione e quale potenziale fattore di rischio per l'incolumità dell'analogico. Un nodo problematico, che rischia di portare i settori culturali tradizionali a un inedito contrasto con il digitale, proprio nel momento in cui le istanze comunitarie lo definiscono invece come parte integrante del Patrimonio Culturale.

Una soluzione efficace a questo problema non può nascere se non nel campo della formazione, agendo con gli strumenti e le strategie innovative delle DH per strutturare una *Catena del Valore Formativo* che doti le nuove generazioni della conoscenza non solo delle nuove dimensioni del Patrimonio culturale, ma del più vasto universo del nuovo DCH.

È, infatti, sicuramente nel campo della formazione che si potrà realizzare l'integrazione fra saperi umanistici tradizionali e conoscenze di metodi e tecniche computazionali nella strutturazione della nuova *Cultura Digitale*, passando dal modello attualmente vigente, finalizzato a creare competenze *borderline* o esclusivamente tecniche, a un nuovo modello che punti a creare conoscenze e competenze consapevoli trasversali, abilitate a partecipare attivamente ai processi di innovazione digitale.

Alcune esperienze didattiche e formative (che non si descriveranno qui per brevità di spazio) sui processi di digitalizzazione del patrimonio culturale e di creazione di contenuti culturali digitali hanno ben mostrato il ruolo fondamentale che rivestirebbero le DH come settore autonomo transdisciplinare rispetto ai settori tradizionali.

## **LE DH NELLA CATENA DEL VALORE FORMATIVO**

Presupposto a nostro parere necessario per collocare le DH nel processo di costruzione di una *Catena del Valore Formativo* è innanzitutto invertire e storicizzare quali tappe ormai superate del processo evolutivo gli approcci al digitale ancora oggi ampiamente utilizzati, che lo

identificano eminentemente in mero strumento *social*, o di semplificazione della gestione quotidiana, o di surroga emozionale utile a fruizione del patrimonio culturale analogico assolutamente estemporanea e vincolata all'*hic et nunc*, senza alcuna prospettiva di sostenibilità dei processi e delle risorse digitali, che invece recenti studi hanno mostrato essere possibile.

È, questa, una sfida complessa, che implica necessariamente il contestuale ripensamento e la riprogrammazione di metodi e tecniche della formazione, riconoscendo le DH come settore autonomo nella formazione *Long Life Learning* (L3) e valorizzandone la funzione di disciplina di riferimento da collocare con un ruolo ben definito nella costruzione dei nuovi saperi, che sviluppi nelle nuove generazioni un approccio al digitale fondato su conoscenze e competenze culturali anche metodologicamente ben strutturate, e le abiliti a riconoscere e validare nei processi di digitalizzazione e di co-creazione digitale, ancorché nelle risorse digitali prodotte, il valore di memoria e di fonte di conoscenza della società contemporanea da conservare e valorizzare nel futuro.

## **DEFINIRE IL DCH, RICONOSCERE LE DH**

Un fattore preliminare necessario per questa innovazione e riprogrammazione è la definizione del concetto di DCH. In un altro nostro contributo pubblicato precedentemente, abbiamo proposto che il riconoscimento del DCH si debba basare sulla validazione e certificazione dei metadati partendo da una implementazione dei *FAIR Principles*. In particolare, abbiamo articolato la *R* di *Re-usable* in *R4* aggiungendo i concetti *Relevant*, *Reliable* e *Resilient*. Partendo da questo assunto, abbiamo delineato una classificazione nelle tre macro categorie *Born Digital Heritage*, *Digital FOR Cultural Heritage*, *Digital AS Cultural Heritage*, dalle quali abbiamo poi inferito la seguente proposta di definizione concettuale del DCH:

Il Digital Cultural Heritage è l'ecosistema di processi, entità, fenomeni Born Digital e Digitized che, essendo stati certificati e validati come conformi ai requisiti FAIR4, fin dalla loro genesi siano potenziali testimonianze, manifestazioni ed espressioni dei processi evolutivi che identificano e connotano ciascuna comunità, contesto socio culturale, ecosistema semplice o complesso dell'Era Digitale, assumendo la funzione di memoria e fonte di conoscenza per le generazioni future.

I concetti chiave che caratterizzano e accomunano le tre categorie e che ritroviamo nella definizione complessiva del DCH sono *processo* e *memoria*: questi, che rappresentano la sostanza sottesa fin dal principio ai ventennali percorsi di ricerca, formazione e didattica delle DH, sono i presupposti che sostanziano l'inderogabilità del riconoscimento del loro ruolo quale settore scientifico disciplinare autonomo indispensabile per la generare *Cultura Digitale* in modalità L3.

Assumendo le definizioni del DCH sopra richiamate, risulta dunque difficile continuare a relegare le DH a mosaico indefinito di risorse strumentali all'innovazione dei saperi tradizionali, negando il ruolo fondamentale che esse invece rivestirebbero nello sviluppo della nuova *Cultura Digitale*. Questa, infatti, è l'obiettivo degli approcci culturali integrati sperimentati nei percorsi didattico formativi già da qualche anno intrapresi nell'ambito delle DH, i cui risultati hanno evidenziato la funzione essenziale del settore nella creazione dell'ecosistema di conoscenze e competenze digitali consapevoli indispensabili alla gestione sia dell'*Era Digitale* contemporanea, sia della sua evoluzione futura.

## **CONCLUSIONI**

Le recenti *Conclusioni* dell'UE sul Patrimonio Culturale, con il riconoscimento del DCH quale suo elemento integrante, pongono la strutturazione di una sostanziale *Cultura Digitale* nelle nuove generazioni al centro degli obiettivi strategici da raggiungere nel più breve tempo



possibile. Il riconoscimento delle DH quale settore scientifico disciplinare autonomo e la loro adeguata collocazione nell'ambito della formazione L3 è quindi diventato tema urgente e primario, in quanto permette di raggiungere importanti obiettivi, che vanno ben oltre l'apprendimento di tecniche e metodologie per la produzione di contenuti e *asset* digitali. Le DH, infatti, consentono di strutturare la *Catena del Valore Formativo* necessaria a formare le conoscenze e competenze interdisciplinari e le sinergie operative tra diversi attori, a guidare i processi di co-progettazione e co-creazione come superamento dell'interdisciplinarietà, a generare consapevolezza dell'intero processo di ricerca/tutela/valorizzazione del Patrimonio Culturale nella sua nuova dimensione. La vera sfida delle DH per il futuro, dunque, consiste nel conferire una nuova sostanza alla disciplina attraverso la valorizzazione della sua dimensione più profonda, liberandosi dei preconcetti e dei pregiudizi dilaganti: è in questa luce che si può iniziare a strutturare, anche e soprattutto in una prospettiva digitale, la Cultura necessaria a costruire il sistema di conoscenze e competenze consapevoli in grado di gestire l'*Era Digitale* in corso e nella sua evoluzione.

## BIBLIOGRAFIA

- [1] N. Barbuti (2016), *Le nuove entità culturali digitali tra Intangible Cultural Heritage e Patrimonio Culturale Immateriale*, in The Creative Network – Conferenza GARR, Firenze, 30 novembre-02 dicembre 2016, <https://www.eventi.garr.it/it/conf16/home/materiali-conferenza-2016/paper/19-conf2016-paper-barbuti/file>
- [2] N. Barbuti (2017), *Dalla Digital Culture al Digital Cultural Heritage: l'evoluzione impossibile?*, in AIUCD 2017 Conference – Book of Abstract. Il telescopio inverso: big data e distant reading nelle discipline umanistiche, p. 14-17, AIUCD <<http://aiucd2017.aiucd.it/wp-content/uploads/2017/01/book-of-abstract-AIUCD-2017.pdf>>.
- [3] N. Barbuti (2018), *From Digital Cultural Heritage to Digital Culture: Evolution in Digital Humanities*, in DTUC '18 Proceeding of the 1st International Conference on Digital Tools & Uses Congress, ISBN: 978-1-4503-6451-5, doi>10.1145/3240117.3240142.
- [4] N. Barbuti, L. Marinucci (2018), *Dal Digital Cultural Heritage alla Digital Culture. Evoluzioni nelle Digital Humanities*, DH 2018, <https://dh2018.adho.org/dal-digital-cultural-heritage-alla-digital-culture-evoluzioni-nelle-digital-humanities/>
- [5] F. Ciotti (a cura di, 2014), *Digital Humanities: progetti italiani ed esperienze di convergenza multidisciplinare. Atti del convegno annual dell'Associazione per l'Informatica Umanistica e la Cultura Digitale (AIUCD)*, Firenze, 13-14 dicembre 2012, Studi Umanistici, Serie Quaderni DigiLab, Sapienza Università Editrice.
- [6] F. Ciotti (2018), *Dall'Informatica Umanistica alle Digital Humanities. Per una storia concettuale delle DH in Italia*, DH 2018, <https://dh2018.adho.org/dallinformatica-umanistica-alle-digital-humanities-per-una-storia-concettuale-delle-dh-in-italia/>
- [7] D. Manacorda (2017), *Patrimonio culturale: un patrimonio collettivo*, in R. Auriemma (a cura di), *La democrazia della conoscenza: patrimoni culturali, sistemi informativi e open data: accesso libero ai beni comuni? Atti del Convegno (Trieste 28-29 gennaio 2016)*, Udine, Forum Edizioni.
- [8] P. Mastandrea (2018), *L'orizzonte delle Digital Humanities*, Umanistica Digitale n. 2, ISSN: 2532-8816 - n.2, 2018, DOI: <https://doi.org/10.6092/issn.2532-8816/8055>.
- [9] P. Monella (2015), *L'Informatica Umanistica tra istituzionalizzazione e strumentalismo*, in *Annuario delle attività 2012 - Contributi del Centro Linceo Interdisciplinare Beniamino Segre*, Accademia Nazionale dei Lincei, ISBN 978-88-218-1705-0.
- [10] J. Schnapp (s.d.), *Di cosa parliamo quando parliamo di Digital Humanities (intervista a Jeffrey Schnapp)*, in *Digital Humanities: il mondo digitale e le città (il*

progetto I-Media-Cities), Cineca Magazine 5\_Human-Technologies DOI 10.1388/magazine-05-03

[11] F. Tomasi (2015), *Discipline umanistiche e informatica. Quale futuro per l'integrazione?*, in *Labor & Law Issues*, vol. 1, n. 1. [12] G. Volpe (2017), *Un'eredità da riconquistare. Libero accesso e riforma dei beni culturali*, in R. Auriemma (a cura di), *La democrazia della conoscenza: patrimoni culturali, sistemi informativi e open data: accesso libero ai beni comuni?* Atti del convegno (Trieste 28-29 gennaio 2016), Udine, Forum Edizioni.

[13] G. Volpe, G. De Felice (2014), *Comunicazione e progetto culturale, archeologia e società*, in "European Journal of Post-Classical Archaeologies" 4. ISSN 2039-7895.

# Creating Digital Culture by digitizing Cultural Heritage: the Crowddreaming living lab method.

N. Barbuti<sup>1</sup>, Giuliano De Felice<sup>2</sup>, Annalisa Di Zanni<sup>3</sup>, P. Russo<sup>4</sup>, A. Valentini<sup>5</sup>

<sup>1</sup> Department of Humanities (DISUM), University of Bari Aldo Moro, Bari, Italy, [nicola.barbuti@uniba.it](mailto:nicola.barbuti@uniba.it)

<sup>2</sup> Liceo Statale Classico, Linguistico e Scienze Umane "F. De Sanctis", Bari, Italy, [annalisa.dizanni@unifg.it](mailto:annalisa.dizanni@unifg.it)

<sup>3</sup> Stati Generali dell'Innovazione, Roma, Italy, [paolo.russo@statigeneralinnovazione.it](mailto:paolo.russo@statigeneralinnovazione.it)

<sup>4</sup> Egina S.r.l., Foligno, Italy, [altheovalentini@egina.eu](mailto:altheovalentini@egina.eu)

## ABSTRACT

Since 2015, the Digital Cultural Heritage, Arts & Humanities School (DiCultHer) – an interdisciplinary network of over 70 Italian organizations including universities, research entities, cultural institutions and associations – has tested models to build soft skills required to co-create, manage, preserve and safeguard digital cultural heritage. This paper outlines the current state of advancement in the development of an innovative living lab methodology named *The Art of Crowddreaming*. Such methodology has been implemented within the network activities, and it has been proven to be able to engage innovators, researchers, schools of any grade and other societal actors as a community in the challenge to invent, co-design and build prototypes of cross-generational digital monuments. The methodology is illustrated by means of two case studies. *Quintana 4D* engaged schools of any grade in the City of Foligno in an trans-disciplinary effort to design, expand and manage the Museater of the Joust of Quintana. *Heritellers* engaged students of “F. De Sanctis” high school for classical studies in the City of Trani in the making of "CastleTrApp", a digital storytelling app and a Museater performance about the famous Swabian Castle of their City.

## KEYWORDS

Crowddreaming, Living Lab, Digitization, Digital Culture, Digital Cultural Heritage, Theory

## INTRODUCTION

The basic principles of what would have become *The Art of Crowddreaming* living lab method (Chesbrough 2003) (Almirall and Wareham 2011) (Pallot 2009) (Travaglini and Sabella 2017) worked successfully for the first time in 2014 in the United States for the launch of the *Treasures & Tales* art exhibition in Wilmington, DE as a local community effort (Guardia di Finanza Italiana 2014).

In 2015 such principles began to be tested in Italy too within the Digital Cultural Heritage, Arts & Humanities School – DiCultHer<sup>1</sup>, an interdisciplinary network of over 70 Italian organizations including universities, research entities, cultural institutions and associations, as the guidelines to plan the contest for schools *Crowddreaming: youth co-create digital culture*. Since then, three editions have been run, reaching out to over 5,000 students and leading to the creation of a small community of teachers<sup>2</sup>.

Given the positive feedback coming from the contest, several members of the DiCultHer network decided to run a recursive prototyping effort in order to transform a successful practice into a living lab methodology, largely inspired by the principles of MITs' *Theory U* (Scharmer 2000) (Scharmer 2018).

---

<sup>1</sup> <https://www.diculther.it/>

<sup>2</sup> <https://www.diculther.it/crowddreaming2018/>

Starting from this state-of-the-art, the researchers of DiCultHer co-designed and developed the innovative living lab methodology *The Art of Crowddreaming*, that has been proven to be able to engage innovators, researchers, schools of any grade and other societal actors as a community in the challenge to invent, co-design and build prototypes of cross-generational digital monuments. Such a challenge fosters the activation of several key competences described in the lifelong learning European framework: digital competence; personal, social and learning competence; Civic competence; Cultural awareness and expression competence. It addresses the RRI's public engagement key theme too. The case studies presented in this paper are the first iterations of the prototyping process.

The methodology is in perpetual development by design as it provides tentative answers to three ever-changing Digital Age challenges to teaching and developing cultural heritage and heritage-related skills.

## THE THREE CHALLENGES

The *Challenge of Speed* is common to most aspects of technology-driven society since XX century. Change is so quick and relentless that often it is no longer possible to weigh-in on extrapolations of past patterns to analyze trends in the course of planning for an uncertain future. Therefore, new social technologies develop to let people learn from the future as it emerges. *Theory U* has proven itself a very effective tool. This theory proposes to face the *Challenge of Speed* through a recursive prototyping process, based on a phase of deep listening of the needs of a social ecosystem, the understanding of one's own role in the change process and then the development of a prototype to drive the change. Nonetheless, *Theory U* is biased towards leading communities or enterprises, so it is not very easy for teachers and students to connect with its language and tools. This challenge requires finding a way to make the very effective *Theory U* concepts and tools accessible to educators, cultural operators and kids.

The *Challenge of Form* is relevant to digital societies. Humankind learned to shape energy into meaningful information only a few decades ago. It is a first in history and the extent of this revolution has yet to be explored and understood. Digital technologies advance blazingly fast and new "energy-shaping" skills to be learned emerge at a very hard-to-keep-with pace. Even the concept itself of "digital" has yet to be clarified. Very often it is still regarded as a dimension beyond the physical plane, where things happens almost by magic. This challenge requires finding a way to help teachers, cultural operators and kids to acquire the right digital mindset and skills in a sustainable way.

The *Challenge of Persistency* is specifically relevant to digital cultural heritage. A construct with tangible mass tends to be permanent and persistent in time. Moreover, it is well known how to preserve it. Energy-based digital constructs are very volatile. Millennials are the first generations called to face the Challenge of Persistency, because their culture is and will be more and more digital. This challenge requires finding a way to help them to pass their digital cultural heritage to the next generations. It requires also finding solutions to transmit digital knowledge, practices, art through the centuries.

## THE ART OF CROWDDREAMING LIVING LAB METHODOLOGY

*The Art of Crowddreaming* living lab methodology relies upon the awareness that any kind of mind is intelligent if it is capable of dreaming. This is true for the connective minds of digital age too. *The Art of Crowddreaming* is the discipline that trains a connective intelligence to lucid dream. A *crowddream* comes to life usually as an individual **insight** that becomes a clear **intent** by way of interacting with one's social circles. It becomes then a compelling **story** about a desirable future that is able to capture the imagination of a big enough crowd so

to make it happen. Finally, it evolves into a well-designed innovation **project** that can shape the shared dream into reality.

There are three critical transitions in the *crowddreaming* process: from insight to intent, from intent to compelling story and from compelling story to responsible innovation project.

The first transition corresponds to the deep listening phase on the left side of a *U* procedure and it can be facilitated with the tools and methods suggested by *Theory U*.

The second and third transitions are placed in the right wing of the *U*, when a prototype has to be created. *The Art of Crowddreaming* educates to manage it by using methods and language of Hollywood blockbuster productions instead of the academic leadership-oriented vocabulary of *Theory U*. Such narrative choice makes the concept more accessible to everyone and the prototyping phase way more appealing to students: they are fascinated by the idea of learning how to produce a blockbuster movie.

The starting point is a quite obvious statement: every project that achieved its goal has a success story to tell. Proceeding backwards from the socially desirable future of the happy ending, innovators, researchers and societal actors co-create the plot with all of its characters, relationships, places, props (resources), events, subplots that had to be there.

Dramatic theory and movie production management knowledge provide highly professional and perfectly honed tools to develop a story from the original insight to its final staging in the real world as required by the Crowddreaming process. They not only force to identify all the necessary human, financial, material and time resources, but they oblige to explore also the emotional, ethical and human dimension, which is often overlooked with purely analytical and numerical approaches to project management.

The "Crowddreaming a Digital Monument" is a framework that facilitates the set up of an art-of-crowddreaming-based living lab experience focused on the development of the soft skills required to co-create, manage, preserve and safeguard digital cultural heritage. The monument is a *museater*<sup>3</sup>, a place where to preserve digital memories about cultural exchanges and at the same time a stage where people are encouraged to act and interact.

Based on the constructivist approach of project-based learning, teachers, students and societal actors are challenged to *crowddream* a digital story about a relevant topic regarding their cultural identity. The story has to be able to travel in time through many future generations. The overall goal is to encourage both teachers and young people to get ready to face the epochal challenge to which the new generations are called: they are the first ones in the history of humankind to find themselves passing down a purely digital cultural heritage.

## THE QUINTANA 4D LAB SUCCESS STORY

The *Quintana 4D Museater Lab* (Q4D) revolves around the 70-year old *Giostra della Quintana* of the city of Foligno. The *Quintana* was born in 1946 as an ahead of its time social innovation experiment and worked so well that it became a tradition.

In the past, kids and teenagers loved to be part of the celebration. Today, young people are leaving the Quintana community at an unprecedented rate. The radically different nature of the Millennial's digital culture is the main reason of such disaffection: interviews showed that young people feel not the Quintana as an expression of their own digital culture.

The *Q4D Lab* aimed at transforming digitization from a menace into an opportunity to engage young people into the events in new and different ways, by involving them in *Museater* activities placed at the crossroad between digital technologies and local fascinating and rich cultural heritage.

---

<sup>3</sup> <https://technical.ly/delaware/2014/10/09/special-exhibit-wilmington-offers-italian-art-seen-apps/>

A *museater* is an entity that cares for collections of scientific, artistic, cultural or historical stories, helps people to experience them in contexts where their educational and emotional impact is maximized, and facilitates their re-use and diffusion: it is a hybridization between museum, theater and digital research lab using AR technologies. It reveals and makes accessible a still empty digital space, which waits for someone to transform it into a place where humankind can live part of its life.

A group of Millennials has been engaged to take this empty digital space over and to make it as beautiful and significant. They were called to *crowddream* and build an invisible city atop of the visible one. Brainstorming meetings guided by a facilitator let them create the success story of their “dream” *museater* about Quintana. The project found its catalyst in a young woman who is now in charge of the *Q4D Museater Lab*. She quite literally dreamed and designed her own new (and first) digital-cultural-heritage-based job. She acted as tutor for apprenticeship programs with high schools, based upon project-based learning where students helped professionals in the process of planning the production of their success story and making the Lab a reality.

Today the *Q4D Museater Lab* is hosted at the beautiful palace Brunetti-Candiotti, a committed and ever-growing group of over 600 Millennials have been working as volunteers to haunt Foligno's streets and buildings with geo-located digital *ghosts*. Moreover, the AR-enhanced experimental exhibition has become both the focus for educational activities of almost every school in Foligno and a destination for occasional visitors and guided tours with over 1,500 visitors in the first 9 months of its life. It is a small win-win ecosystem already, where young people feel that they do not need to find their place in society: they can build it by exploring and shaping the digital dimension of their city.

## **THE CASTLE TRAPP SUCCESS STORY**

The *Heritellers Lab* comes from the awareness of the engagement potential of the Swabian Castle of Trani, one of the most important monuments built by Emperor Federicus II.

In this project, the students of “G. De Sanctis” high school of Trani realized another way of *Museater Lab* by undertaking a number of initiatives (animated visits, recreational-didactic activities, creation of digital audio-video storytelling and swipe stories, etc.) that make the school a civic center, where they design together with- and for the territory activities that develop citizenship in the name of an identity inclusive idea. By this way, the goal of the project has been to promote a sense of cultural ownership and to provide students with a global vision of all aspects of cultural heritage (research, protection, management, use, production), fostering an approach oriented to communication and enhancement and opening up to the use of digital technologies for promoting and sharing cultural heritage.

Students were challenged to imagine the co-creation of a successful app to tell the Federicus II cultural innovation, guiding visitors to discover the true significance of the Swabian Castle of Trani. They were asked to plan all the activities required to create the success story of their *museater*. Following the *living lab* methodology, they were asked to test their assumptions by undertaking the initiatives above mentioned.

The first year activities aimed at creating a *swipe story* about the Swabian Castle: an innovative digital multimedia storytelling which develops on an illustrated tape, animated and interactive on multiple levels, and uses a simple and immediate language based on drawings, images, words, games, sounds and movies. By this storytelling, a process of *cultural ownership* starts that makes both students aware and responsible custodians of cultural heritage inherited from the past, and creators of potential digital cultural heritage. The *swipe story* is entitled *CastleTrApp* and it is produced in Italian and English. It tells the most

important historical phases of the Swabian Castle and can be downloaded by Google Play<sup>4</sup> and App Store<sup>5</sup>.

By retracing some of the transformations undergone over the centuries, the app allows the visit of the Castle using tools of digital storytelling. In a simple and fun way, a very large audience can explore the castle personalizing its experience, thanks to the different levels of depth of the proposed content, all scientifically validated.

Consistent with the theme of the project, in 2018 “De Sanctis” students involved about 340 children of “Roncalli” primary school of Altamura, realizing an animated and playful visit of the Castle to discover the monument by reproducing the app storytelling. In the so-called “Sala Federico II” the designers of the app briefly explained the project to children. Other students, as guides, explained to children the mechanism of the game. During the visit, at a given signal the children closed their eyes and, reopening them, they found themselves in a Medioeval historical phase characterized by the “De Sanctis” students in *museater* way. From that moment the children visited the Castle rooms equipped with medieval furniture and populated by some costumed characters also reproduced in *the swipe story* of the app.

*Heritelles* project realized another kind of small win-win ecosystem, where young people feel that they do not have mandatorily to find their place in society since still kids.

## CONCLUSION

This paper outlines the key principles of *The Art of Crowddreaming*, a living lab methodology that was elaborated for the first time in 2014 in the United States. A crowddream comes to life usually as an individual insight that becomes a clear intent by way of interacting with his/her social circles. It becomes then a compelling story about a desirable future that is able to capture the imagination of a big enough crowd so to make it happen. Finally, it evolves into a well-designed innovation project that can shape the shared dream into reality. This methodology has been proven to be able to engage innovators, researchers, schools of any grade and other societal actors as a community in the challenge to invent, co-design and build prototypes of cross-generational digital monuments.

*The Art of Crowddreaming* is strongly oriented towards public engagement and it is designed to stimulate the development of some of the key competences described in the lifelong learning European framework: digital competence; personal, social and learning competence; Civic competence; Cultural awareness and expression competence. Its innovative component resides mainly in its intent to insure a broader impact of *Theory U* by making its key practices and concepts more accessible to educators, youth workers, teenagers and kids by using familiar and engaging tools borrowed by movie’s industry rather than the original quite philosophical and academic language.

The experiments we tested show that this methodology works and it is adaptable to very different scales: from a single school class to a whole community.

## BIBLIOGRAPHY

- [1] Alcotra Innovazione (2013), *La creazione di Living Lab Transfrontalieri*, <http://www.alcotra-innovazione.eu/livingLabs/dwd/Manuale%20Alcotra%20Innovazione.pdf>
- [2] Chesbrough, H.W. (2003), *Open Innovation: The new imperative for creating and profiting from technology*. Boston: Harvard Business School Press.

---

<sup>4</sup> <https://play.google.com/store/apps/details?id=com.ia2.castleTrAPP>

<sup>5</sup> <https://itunes.apple.com/it/app/castletrapp/id1296811891?mt=8>

- [3] Guardia di Finanza (2014), *Treasures and Tales of Italy's*, Guardia di Finanza Art Recovery Team DeBooks.
- [4] Pallot M. (2009), *Engaging Users into Research and Innovation: The Living Lab Approach as a User Centred Open Innovation Ecosystem*. Webergence Blog. Copia archiviata, su cwe-projects.eu. URL consultato il 7 giugno 2011 (archiviato dall'url originale il 9 maggio 2012).
- [5] Scharmer, C.O. (2000), *Presencing: Learning From the Future As It Emerges Presented at the Conference On Knowledge and Innovation*, May 25-26, 2000, Helsinki School of Economics, Finland, and the MIT Sloan School of Management, OSG, October 20th, 2000.
- [6] Scharmer, C.O. (2018), *The Essential of Theory U: Core Principles and Applications*, Berrett-Koehler Publishers.
- [7] Ståhlbröst, M.H. (2012), *The Living Lab Methodology Handbook*, Centre for Distance-spanning Technology, Sweden
- [8] Almirall, E., Wareham, J. (2011), *Living Labs: Arbiters of Mid- and Ground- Level Innovation*, in *Technology Analysis and Strategic Management*, 23(1), 2011 pp. 87-102.
- [9] Travaglini, F., Sabella, P. (2016), *Scuola 4.0, living labs: le scuole che progettano il futuro*, in <http://www.zonalocale.it/2016/11/30/scuola-40-living-labs-le-scuole-che-progettano-il-futuro/23394?e=sansalvo>
- [10] <https://www.diculther.it/>
- [11] <https://www.diculther.it/crowddreaming2018/>
- [12] <https://technical.ly/delaware/2014/10/09/special-exhibit-wilmington-offers-italian-art-seen-apps/>
- [13] Audiovisual Living Lab Terrassa [Spagna]: [www.alivinglab.com/alten](http://www.alivinglab.com/alten)
- [14] Design Creative Living Lab [Francia]: [www.citedudesign.com/fr/home/](http://www.citedudesign.com/fr/home/)
- [15] [European Network of Living Labs: www.openlivinglabs.eu/](http://www.openlivinglabs.eu/)
- [16] InnovaLab [Spagna]: [blog.innovalab.org/](http://blog.innovalab.org/)



# On the Use of Terminological Records in Specialised Translation

Federica Vezzani <sup>1</sup>, Giorgio Maria Di Nunzio <sup>2</sup>

<sup>1</sup> Department of Linguistic and Literary Studies, University of Padua, [federica.vezzani@phd.unipd.it](mailto:federica.vezzani@phd.unipd.it)

<sup>2</sup> Department of Information Engineering, University of Padua, [giorgiomaria.dinunzio@unipd.it](mailto:giorgiomaria.dinunzio@unipd.it)

## ABSTRACT

In this paper, we focus on the teaching of specialised translation and, in particular, on the preliminary phase of the translation process which is based on a broad and systematic work on the terminology of the micro-language considered. We present a new model of bilingual terminological record, as a digital tool supporting the process of translation of medical documents. Finally, we describe the results of a set of experiments which we have run since 2017 with two groups of students of the master's degree of the University of Padua.

## PAROLE CHIAVE

Terminology, Specialised Translation, Terminological Record, Termbase.

## INTRODUCTION

In recent years, the need for interdisciplinarity has become increasingly important in the field of teaching and it represents a constant goal in all modern educational theories [1]. In this context, the use of digital tools for teaching Humanities has proved to be effective and efficient. For instance, the use of systems that automatically support the process of translation of documents [2] has become widespread in the teaching of Translation Studies [3, 4]. In particular, the digital tools that help to teach the methodology behind the translation of specialised texts [5, 6] facilitate students in understanding the complex translation process. Computer-assisted translation (CAT) tools, such as SDL Trados,<sup>1</sup> Omega T,<sup>2</sup> MateCat,<sup>3</sup> etc., make the translation of a specialised text an interactive process between human and computer [7].

In this paper, we focus on the preliminary phase of the translation process which is based on a broad and systematic work on the terminology of the specialised language considered [8–10] such as medical, economic or legal language. Indeed, the study of terminology and the use of digital terminological resources are fundamental for students in order to 1) create their basis of knowledge in specialised disciplines, 2) master the terminology of a specific micro-language and 3) collect all the relevant linguistic data for both the source and target languages involved.

The collection and organisation of terminological data are performed through terminological records [11] providing linguistic information both for the term in the source language and for the translating candidate. These records are usually gathered into specific terminological databases which constitute a valid support for translators, terminologists and language professionals in general [12–14].

In this paper, we illustrate a new model of bilingual terminological record for the translation of medical documents which has been used in two courses of the master's degree in Modern

---

<sup>1</sup> <https://www.sdltrados.com>

<sup>2</sup> <http://omegat.org>

<sup>3</sup> <https://www.matecat.com>

Languages for the International Communication and Cooperation of the University of Padua: Computer Assisted Translation Tools and Specialised Translation. This digital terminological record is based on the TriMED model proposed in [15] and, in this work, we aim to test its usefulness both in the terminological analysis and in the translation process of such specialised texts.

This paper is organised as follows: in Section 2, we contextualise the process of translation based on a terminological approach; then, we describe our terminological record as a digital tool for teaching specialised translation. In Section 3, we present the experiments with the students. Finally, in Section 4 we give our conclusions and some hints on future works.

## TEACHING TERMINOLOGY FOR MEDICAL TRANSLATION

Technical-scientific translation concerns the process of transmission of technical information from a source language into a target language in specialised domains of human activity [16]. The main difficulties in learning, in particular, the medical translation process are the lack of medical knowledge of students and their unfamiliarity with medical terminology and phraseology [17, 18]. Amal Jammal states that “[...] *on ne peut traduire que ce que lon comprend bien*”, that is to say it is necessary to understand and familiarize with the topics covered in the specialised text before starting the translation process [19].

The screenshot shows the TriMED interface with a dark sidebar on the left containing 'Patient', 'Translator', and 'Physician'. The main area is divided into two columns. The left column is for the source language (English) and the right for the target language (French). Each column has a 'Translator' header and four sections: 'Technical Term', 'Informative Term', 'Popular Term', and 'Definition'. Below these is an 'Analysis' section with a phonetic transcription. The English side shows 'Ecchymosis' as the technical term, 'Bruise/Hematoma' as the informative term, 'Bruise' as the popular term, and a detailed definition. The French side shows 'Ecchymose' as the technical term, 'Ecchymose' as the vulgarisateur, 'Bleu' as the popular term, and a definition in French. The analysis for English is '/skin//spot//bleeding/' and for French is '/Extravasation//sang//tissu//organes//peau/'.

Figure 1: Part of a TriMED terminological record for the technical term in the source language (EN) Ecchymosis and its equivalent in the target language (FR) Ecchymose.

Therefore, the translation process is usually divided into two main operations: the decoding and the subsequent transcoding of medical information [19, 20]. The first step consists in the decoding of technical terms conveying medical information in the source language. This means that, in order to understand specialised notions, students (as future professional

translators) should create their terminological corpus, both in the source and target languages, gathering most of the relevant documents for the topic to be translated and becoming familiar with medical terminology. The second phase, involving the process of translation itself, concerns the transcoding of the lexical units previously identified into the target language by respecting the faithful transmission of information from a semantic viewpoint and the morphological and syntactic structure of the working language. In this context, terminological records are commonly used as tools gathering information about a term, its translating candidate and the related concept [11, 21].

## TriMED TERMINOLOGICAL RECORD

The new model we propose is TriMED terminological record which was designed for language professionals and it has been used by students while learning the terminological analysis and the translation process itself [15]. This resource provides all the information necessary in order to decode and then transcode the meaning of a technical term and its correct translating candidate. It is structured around four axes of analysis:

1. Formal features;
2. Semantics;
3. Corpus;
4. References.

Regarding the formal and lexical framework of the term, our model of terminological record provides information such as: gender, spelling, pronunciation in the International Phonetic Alphabet (IPA) and other information about the etymology, such as derivation and composition of the term. Based on the WordNet resource,<sup>4</sup> the record contains also all the nouns, verbs, adjectives, and adverbs deriving from the analysed term and which fall into the same semantic sphere.

The second section focuses on the semantic features of the term. First, we propose a definition extracted from reliable resources such as Merriam-Webster Medical Dictionary<sup>5</sup> or MediLexicon<sup>6</sup> especially for acronyms and abbreviations. In addition, we provide the semic analysis of the term [22] that is a methodology used in compositional semantics in order to decompose the meaning of technical terms (lexematic or morphological unity) into minimal units of meaning: the semes. Moreover, in order to evaluate the semantic behaviour, we collect the phraseology of the term by considering cases of collocations [23] and colligations [24].

In the corpus section, we provide all specialised contexts where technical terms have been extracted and then we proceed through the identification of the domain and the register of communication of the term (popular, standard and specialised). The term and its definition, therefore, take on meaning when they are connected to a specific domain: in our analysis, we identify the domain and subdomains of the text (such as surgery, pathology, pharmacology, etc). Finally, since all of this information has been extracted from different sources, we provide references to each source.

The Web application that allows users to manage and store TriMED terminological records was realised with the Shiny R package [25]. The interface was specifically designed to display the information of each field in the source and target language, as shown in Figure 1. Currently, the TriMED terminological database supports three working languages (English,

---

<sup>4</sup> <https://wordnet.princeton.edu>

<sup>5</sup> <https://www.merriam-webster.com/>

<sup>6</sup> <https://www.medilexicon.com/>

French and Italian) and contains about 328 terminological records that are under review. A demo of the application is available.<sup>7</sup>

## EXPERIMENTS

In this section, we present the studies that we have carried out on the use of the TriMED terminological record for teaching the translation process of medical texts both in the decoding and transcoding phase. In particular, we describe the educational applications concerning the use of terminological records by students of the master's degree course in Modern Languages for the International Communication and Cooperation of the University of Padua.

### DECODING PHASE

The decoding phase of a specialised translation concerns the collection of a corpus of texts that can form the basis of preliminary knowledge in the process of translation of texts. In 2017/18, we performed an experiment with 90 students of the course on Computer Assisted Translation Tools. The objective of this experiments was to teach: i) how to correctly interpret pieces of medical information about a disease provided by a physician and ii) how to produce some new information that might help the physician to find relevant documents about the disease.<sup>8</sup>

Students were divided in groups of 3 people. We provided each group with a specialised medical query about a disease written by a physician and a scientific article (related to the query) retrieved from the PubMed database<sup>9</sup> (one of the largest databases of biomedical literature). Each group was asked to rewrite the original query by following a terminological approach and produce an alternative that might have helped the physician to find additional relevant documents about the disease. Finally, they were asked to translate the scientific article provided. In order to complete the task, we asked to the students to extract all the technical terms and to fill-in the TriMED terminological records. In this way, the students were able to collect all the terminological data related to the extracted terms and become familiar with the medical field before starting the translation process. The results of the experiment showed that, by means of the TriMED record, the students were able to decode correctly the specialised text and produce some new medical information that had the same quality of the original pieces of information provided by the expert [27]. For example:

- Original Query: PET-CT for assessing mediastinal lymph node involvement in patients with suspected resectable non-small cell lung cancer
- Alternative Query: Positron emission tomography Computed tomography for the evaluation of mediastinum lymph nodes in doubtful cancer affected people with removal NSCLC.

At present, we are running a second experiment with the students of the course of 2018/19.

### TRANSCODING PHASE

The second step of the translation process concerns the transcoding of the information extracted from the text in the source language into the target language in a clear and faithful

---

<sup>7</sup><https://gmdn.shinyapps.io/TriMED/>

<sup>8</sup>This study was part of a larger experiment that involved our participation to the Cross-Language Evaluation Forum (CLEF) task on Technologically Assisted Re-views [26]. For space reasons, we omit the details of the technical part.

<sup>9</sup><https://www.ncbi.nlm.nih.gov/pubmed/>

way. Currently, we are performing an experiment concerning the use of TriMED terminological records for teaching Specialised Translation in the medical domain to the students of the second year of the Master's Degree course.

Students work with Italian and French languages and they are asked to translate two documents (French to Italian and Italian to French) about the practice of vaccinations. In particular, they are translating two pedagogical dossiers for collective protection extracted from the Santé Publique - France website and the Local Social Health Unit of Padua website. Students are working on the technical terms extracted from the specialised texts, and they are compiling terminological records in order to choose the correct translating candidate.

After the manual extraction of technical terms, they proceed with the creation of a new empty terminological record from the TriMED application and then they manually fill all its fields with linguistic information extracted from reliable web sources. For each terminological record we save the creation date and the author and students have the possibility to modify and adjust the information provided. In order to avoid ambiguity from the semantic viewpoint, students have to fulfil the records with terminological data related only to the specific field or subfield of interest (in our case, medical field and in particular vaccination subfield).

At the moment, terminological records are proving to be a valid support for the students in the transmission of medical information. Moreover, TriMED has a user-friendly interface and this allows quick data entry and consultation.

## CONCLUSIONS

In this paper, we focused on the preliminary phase of the specialised translation process which is based on the study of technical terms and their analysis by means of terminological records. In particular, we presented our methodology for the translation of medical texts through TriMED terminological record which is structured in order to provide a useful support for students (as future translators) in the two major operations involved in the translation process: the decoding and the subsequent transcoding of medical information.

We performed two experiments with the students of the master's degree course of the University of Padua and the model we proposed proved to be valid for the: i) creation of students' basis of knowledge in medical domain, ii) mastery of the terminology and iii) collection of all the relevant linguistic data for both the source and target languages. As future work, we are working on a method in order to import the terminological data obtained through the TriMED application on an assisted translation tool like SDL Trados.

## REFERENCES

- [1] Rossi, P.G., Biondi, S.: Interdisciplinarietà. *Education Sciences & Society* 5(1) (2014)
- [2] Hutchins, W.J.: Machine translation over fifty years. *Histoire épistémologie langage* 23(1) (2001) 7–31
- [3] Toury, G.: *Descriptive Translation Studies and beyond: revised edition*. Volume 100. John Benjamins Publishing (2012)
- [4] Gambier, Y., Van Doorslaer, L.: *Handbook of translation studies*. Volume 1. John Benjamins Publishing (2010)
- [5] Scarpa, F.: *La traduzione specializzata. Un approccio didattico professionale*. Hoepli (2008)
- [6] Musacchio, M.T.: Terminology and specialized translation. *Textus online only*. 12 (1999), N. 2, 1999 12(2) (1999) 1000–1021
- [7] Hutchins, J.: Machine translation and human translation: in competition or in complementation. *International Journal of Translation* 13(1-2) (2001) 5–20

- [8] Cabré, M.T.: Terminology and translation. Gambier, Yves & Luc van Doorslaer (eds.) (2010) 356–365
- [9] Wright, S.E., Budin, G.: Handbook of terminology management: application- oriented terminology management. Volume 2. John Benjamins Publishing (2001)
- [10] L’Homme, M.C.: La terminologie: principes et techniques. Pum (2004)
- [11] Gouadec, D.: Terminologie: constitution des données. AFNOR gestion. AFNOR (1990)
- [12] L’Homme, M.C.: Dicoinfo, dictionnaire fondamental de l’informatique et de l’internet. Consulté à l’adresse <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi> (2009)
- [13] L’Homme, M.C., Lanneville, M.: Dicoenviro. dictionnaire fondamental de l’environnement. Consulté à l’adresse <http://olst.ling.umontreal.ca/cgi-bin/dicoenviro/search.cgi> (2014)
- [14] Johnson, I., Macphail, A.: Late–inter-agency terminology exchange: Development of a single central terminology database for the institutions and agencies of the european union. In: Proceedings of the Workshop on Terminology resources and computation, LREC 2000 Conference. Athènes, Grèce. (2000)
- [15] Vezzani, F., Di Nunzio, G.M., Henrot, G.: Trimed: A multilingual terminological database. In: Proceedings of the Eleventh International Conference on Language Resources and Evaluation LREC 2018, Miyazaky, Japan, May 7-12, 2018.
- [16] Pinchuck, I.: Scientific and technical translation. Westview Press (1977)
- [17] Wakabayashi, J.: Teaching medical translation. *Meta* 41(3) (1996) 356–365
- [18] M., R.: La terminologie médicale et ses problèmes. *Tribuna*, Vol. IV, n. 12 (2003)
- [19] Jammal, A.: Une méthodologie de la traduction médicale. *Meta: journal des traducteurs/Meta: Translators’ Journal* 44(2) (1999) 217–237
- [20] Rouleau, M.: La langue médicale: une langue de spécialité à emprunter le temps d’une traduction. *TTR: traduction, terminologie, rédaction* 8(2) (1995) 29–49
- [21] Francoeur, A.: La fiche terminologique, entre théorie et pratique. *Langues et linguistique* (35) (2015) 24–39
- [22] Rastier, F.: Sémantique interprétative. Formes sémiotiques. Presses universitaires de France (1987)
- [23] Firth, J.R.: A synopsis of linguistic theory, 1930-1955. *Studies in linguistic analysis* (1957)
- [24] Sinclair, J.: Reading concordances: An introduction. Pearson Longman (2003)
- [25] CHANG, W.: Shiny: Web Application Framework for R. (2015) R package version 0.11.
- [26] Di Nunzio, G.M., Beghini, F., Vezzani, F., Henrot, G.: An interactive two- dimensional approach to query aspects rewriting in systematic reviews. IMS unipd at CLEF ehealth task 2. In: Working Notes of CLEF 2017 - Conference and Labs of the Evaluation Forum, Dublin, Ireland, September 11-14, 2017. (2017)
- [27] Di Nunzio, G.M., Vezzani, F.: Using R markdown for replicable experiments in evidence based medicine. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction - 9th International Conference of the CLEF Association, CLEF 2018, Avignon, France, September 10-14, 2018, Proceedings. (2018) 28–39

# EpiCUM. Un museo epigrafico digitale per visitatori e specialisti

Salvatore Cristofaro, Daria Spampinato

Istituto di Scienze e Tecnologie della Cognizione - CNR, Italia {salvatore.cristofaro,  
daria.spampinato}@istc.cnr.it

## ABSTRACT

In questo lavoro viene descritto lo stato di avanzamento del progetto EpiCUM che si propone di presentare e rendere fruibile con un unico museo digitale tutto il corpus epigrafico del museo civico Castello Ursino di Catania, codificato in *EpiDoc*.

Il progetto nasce dalla collaborazione tra l'Istituto di Scienze e Tecnologie della Cognizione del CNR e il Comune di Catania e ha visto l'interesse del progetto I.Sicily ed il coinvolgimento del liceo artistico M.M. Lazzaro di Catania con le attività di alternanza scuola-lavoro. La prima fase del progetto è stata dedicata soprattutto alla ricognizione delle epigrafi e all'allestimento della mostra *Voci di Pietra* che propone una selezione di epigrafi secondo modalità di esposizione innovative attraverso l'uso intelligente della tecnologia e del digitale. Nella fase successiva ci si è concentrati sull'analisi e sulla codifica nel formato aperto EpiDoc di tutte le iscrizioni del museo civico e la realizzazione del museo digitale. Attraverso questa codifica e la messa online dei file si rende fruibile, per diversi scopi, il patrimonio epigrafico del museo civico sia agli studiosi del settore, ma anche a tutte le altre tipologie di possibili utenti.

Tutte le informazioni sulle epigrafi sono contenute nei file EpiDoc XML e vengono interrogate opportunamente, presentando i dati in maniera intuitiva nel front-end. Proprio per facilitare la navigazione attraverso le informazioni epigrafiche, si è scelto di codificare nei file XML sia l'ambito religioso, sia le collezioni di provenienza, nonché lo status di copia (o falso) delle epigrafi utilizzando uno specifico set di elementi di marcatura EpiDoc.

## PAROLE CHIAVE

Epigrafia Digitale, Museo epigrafico, Castello Ursino Catania, EpiDoc, religione, copie

## INTRODUZIONE

Il progetto EpiCUM (*Epigraphs of Castello Ursino Museum*) si propone di codificare in EpiDoc<sup>1</sup> e di presentare e rendere fruibile con un'unica interfaccia web tutto il corpus epigrafico del museo civico Castello Ursino di Catania. Il progetto nasce dalla collaborazione tra l'Istituto di Scienze e Tecnologie della Cognizione del CNR e il Comune di Catania e ha visto l'interesse del progetto I.Sicily<sup>2</sup> ed il coinvolgimento del liceo artistico M.M. Lazzaro di Catania con le attività di alternanza scuola-lavoro.

La prima fase del progetto, descritta in [1], è stata dedicata soprattutto alla ricognizione delle epigrafi e all'allestimento della mostra *Voci di Pietra* che propone una selezione di epigrafi secondo modalità di esposizione innovative, attraverso l'uso intelligente della tecnologia e del

---

<sup>1</sup> EpiDoc (Epigraphic Documents) è un progetto internazionale open source di collaborazione tra umanisti e informatici (<https://sourceforge.net/p/epidoc/wiki/Home/>), adottato dalla comunità internazionale degli epigrafisti, con l'obiettivo di creare degli standard flessibili ma scientificamente rigorosi, e degli strumenti per la codifica digitale.

<sup>2</sup> Progetto che si pone l'obiettivo della catalogazione digitale in EpiDoc dell'intero patrimonio epigrafico della Sicilia antica dal VII sec. a.C. al VII sec. d.C. consultabile all'indirizzo <http://sicily.classics.ox.ac.uk/>.

digitale. La mostra è stata realizzata sviluppando un nuovo punto di vista comunicativo del patrimonio epigrafico, con la ricostruzione stilizzata di un colombario, la restituzione grafica di una tomba a "edicoletta" e tramite l'uso di materiale propriamente scenografico come la ricostruzione in plexiglass di una statuetta di *Venus Victrix*. Innovativi sono stati l'uso di video, veri e propri esempi di *storytelling* per immagini e la presenza di un chiosco multimediale *touch*, che consente di esplorare il materiale esposto e di approfondire i contenuti delle epigrafi.

Nella fase successiva del progetto ci si è concentrati sull'analisi e sulla codifica nel formato aperto EpiDoc di tutte le iscrizioni del museo civico e la realizzazione del museo digitale. Attraverso la codifica EpiDoc delle iscrizioni e la messa online dei corrispondenti file XML si rende fruibile, per diversi scopi, il patrimonio epigrafico del museo civico sia agli studiosi del settore, ma anche a tutte le altre tipologie di possibili utenti.

Tutte le informazioni sulle epigrafi sono quindi contenute nei file XML e vengono interrogate opportunamente, presentando i dati in maniera intuitiva nel front-end. Proprio per facilitare la navigazione attraverso le informazioni epigrafiche, si è scelto di codificare in EpiDoc sia l'ambito religioso, sia le collezioni di provenienza, sia la condizione di copia delle epigrafi<sup>3</sup>.

## CODIFICA DELLE EPIGRAFI SECONDO EPIDOC

Parte dell'attività di codifica delle epigrafi è stata rivolta all'uniformità dei dati raccolti dalle osservazioni autoptiche svolte al museo civico e delle informazioni recuperate dal volume di Korhonen [2], in cui sono studiate e descritte tutte le epigrafi del museo civico catanese, anche alla luce della loro fruizione via web da diverse tipologie di utenti.

Questa analisi ha comportato un ulteriore affinamento nello schema di codifica TEI-XML, in conformità con lo schema EpiDoc, in modo da tenere conto di altri aspetti ritenuti rilevanti alla descrizione delle epigrafi stesse. Come descritto nella successiva sezione 3, le epigrafi vengono presentate secondo diversi percorsi tematici. La fig. 1 mostra una schermata dell'interfaccia per un'epigrafe sepolcrale, la corrispondente codifica in formato EpiDoc XML e l'immagine.

The figure shows a digital interface for an epigraphic record. On the left, there is a snippet of XML code in EpiDoc format. On the right, there is a metadata table with two columns: 'Dettagli' and 'Commento'. Below the table is an image of a stone epigraphic inscription.

Dettagli	Commento
Materiale	◆ Marmo
Tipo di oggetto	◆ Urna cineraria
Dimensioni	◆ Altezza 18 cm, Larghezza 25,5 cm, Profondità 25 cm
Data dell'iscrizione	◆ Seconda metà del I o prima metà del II sec. d.C.
Contenuto dell'iscrizione	◆ 1. Dis - Manib(us) ◆ T(it)o - Iunio - Festo ◆ 3. v(oxit) - a(nnis) - XLIII - Iunia ◆ 4. D(omi)ni - ...

Fig. 1 - Scheda epigrafica.

<sup>3</sup> Si osservi che una consistente parte della collezione epigrafica del museo civico Castello Ursino è composta proprio da falsi e da copie settecentesche di iscrizioni.



Per agevolare la creazione dei file EpiDoc XML delle epigrafi da parte anche di personale non esperto, è stato realizzato un apposito *form* HTML in cui è possibile inserire i dati relativi alle epigrafi in formato *plain text*. Successivamente, mediante l'esecuzione di appropriate funzioni di *parsing* e di trasformazione sintattica realizzate *ad hoc* per il progetto (per lo più implementate nel linguaggio JAVA), a partire dai dati testuali inseriti nel form vengono generati i corrispondenti file XML contenenti la codifica EpiDoc. Si noti che viene convertito in EpiDoc anche il testo delle epigrafi, effettuandone, tra l'altro, le codifiche relative alle condizioni di conservazione e all'interpretazione del testo stesso (lacune, lettere danneggiate, correzioni, integrazioni, abbreviazioni, ...). Tali codifiche vengono realizzate utilizzando gli appositi elementi di marcatura EpiDoc, secondo gli standard proposti e adottati dalla comunità internazionale degli epigrafisti digitali. Così, ad esempio, in accordo alle linee guida EpiDoc,<sup>4</sup> l'elemento di marcatura `<unclear>` viene usato per codificare caratteri del testo incompleti o mal-eseguiti, la cui interpretazione risulterebbe ambigua al di fuori del loro stesso contesto; similmente, l'elemento `<supplied>` viene impiegato per indicare del testo non esplicitamente presente nel sorgente, ma introdotto in maniera congetturale dell'editore con lo scopo di colmare un gap apparente. Tali elementi di marcatura possono essere facilmente introdotti con il form HTML mediante l'uso di appositi *bottoni dedicati* che ne permettono di definire lo *scope* all'interno della codifica del testo stesso.

Molto lavoro di *data entry* è stato svolto per normalizzare le informazioni bibliografiche e per l'identificazione e la classificazione delle immagini fotografiche ritraenti i manufatti epigrafici. Sono inoltre stati codificati e normalizzati i nomi presenti nel testo per permettere la realizzazione automatica degli indici dei nomi stessi. È stata completata la ricognizione di tutte le epigrafi conservate nel museo civico (più di 500 epigrafi, per la maggior parte allocate nei depositi), aggiornando il catalogo inventariale con la collocazione attuale e le foto fronte e retro delle epigrafi.

## PROPOSTE DI NUOVE CODIFICHE

Al fine di facilitare la navigazione attraverso le informazioni epigrafiche, sono state codificate con dei tag appositi nei file XML le informazioni relative all'ambito religioso, alle collezioni di provenienza e alle copie.

- **Ambito religioso.** Negli schemi di codifica dei corpora epigrafici in EpiDoc finora disponibili in letteratura, l'informazione riguardante il contenuto religioso del testo delle iscrizioni viene conservato in maniera implicita all'interno dei testi stessi; al fine di esplicitare tale informazione si è scelto di utilizzare l'elemento di marcatura `<domain>`, con attributo `type="rel"`, collocandolo immediatamente all'interno dell'elemento `<textDesc>`. Così, a esempio, l'informazione che l'ambito religioso del testo di una data epigrafe è cristiano, viene codificata formalmente come segue:

```
<domain type="rel">epigrafe cristiana</domain>
```

- [6] **Collezioni di provenienza.** Per quanto riguarda la codifica dell'informazione circa le collezioni di provenienza delle epigrafi è stato usato l'elemento `<provenance type="transferred" subtype="given">` all'interno dell'elemento `<history>`.
- [7] **Copie.** Per la codifica dello status di copia di un'epigrafe, è stato proposto l'elemento di marcatura TEI `<derivation>` (all'interno di `<textDesc>`), con attributo

<sup>4</sup> <http://www.stoa.org/epidoc/gl/latest/toc-it.html>.

type="copy"; secondo le specifiche della codifica TEI, l'elemento <derivation> consente infatti di descrivere la natura ed il grado di originalità del testo.

Per quello che ci risulta, le scelte indicate sopra per la codifica della religione, delle collezioni di provenienza e per le copie non sono mai state esplicitate ed adottate in letteratura. In particolare, relativamente alle copie, menzioniamo il recente progetto PRIN 2015 «*False testimonianze. Copie, contraffazioni, manipolazioni e abusi del documento epigrafico antico*» che rende accessibile all'indirizzo <http://edf.unive.it> un corposo database di falsi epigrafici e copie. Le schede ricercabili delle epigrafi non dispongono, tuttavia, di link per scaricare/visualizzare file XML, né viene indicata alcun tipo di codifica eventualmente adoperata per rappresentare le epigrafi stesse.

## PROGETTAZIONE E REALIZZAZIONE MUSEO DIGITALE

Al fine di soddisfare l'obiettivo di promuovere online il patrimonio epigrafico del museo civico Castello Ursino di Catania sia agli utenti specialisti, sia ai visitatori occasionali e non esperti del settore sono state realizzate diverse modalità di visualizzazione e ricerca delle iscrizioni.

Il museo digitale è stato progettato come sistema multicanale, utilizzando appropriate tecniche di responsive web design (RWD) e con opportuni fogli di stile e sono stati adottati gli standard HTML 5, CSS 3, Javascript, Angular JS.<sup>5</sup>

Sul lato *front-ent*, ogni pagina del museo digitale presenta un unico layout che caratterizza il museo digitale stesso, con loghi, font, menù, collegamenti ipertestuali ed altri elementi grafici utili alla navigazione e alla fruizione dei contenuti (fig. 2). L'interfaccia del museo digitale è pensata per un pubblico generalista del museo epigrafico: dal semplice curioso, all'appassionato, allo specialista/studioso del settore. Essa si compone di diverse sezioni.



Fig. 2 - Front-end del museo digitale EpiCUM.

**EPIGRAFI** con le sottosezioni ELENCO, BIBLIOGRAFIA e RICERCA.

<sup>5</sup> Il museo digitale è accessibile online attraverso il link <http://epicum.istc.cnr.it/>.

- Tramite **ELENCO** si accede a una lista di tutte le iscrizioni, suddivisa per pagine, con le foto, il testo e i numeri identificativi di ogni singola epigrafe.
- In **BIBLIOGRAFIA** sono presenti i corpora e i repertori classici delle iscrizioni (CIL, IG, ...) e alcuni testi in cui sono analizzate le iscrizioni del museo civico catanese.
- **RICERCA** permette di effettuare ricerche semplici ed avanzate sui diversi campi della scheda: *materiale, tipo di oggetto, tipo di iscrizione, tecnica di esecuzione, luogo di origine, collezione di provenienza, lingua dell'iscrizione, ambito religioso, data dell'iscrizione, bibliografia e testo dell'iscrizione*. È inoltre possibile filtrare i risultati delle ricerche *intersecandoli* tra loro, creando l'insieme *complementare* ed estraendo, ad esempio, tutte e sole le iscrizioni che corrispondono a copie, ovvero le iscrizioni appartenenti a ben determinati periodi storici, mediante l'utilizzo, in quest'ultimo caso, di una *timeline* che raggruppa le epigrafi in intervalli temporali di 1 secolo (fig. 3).

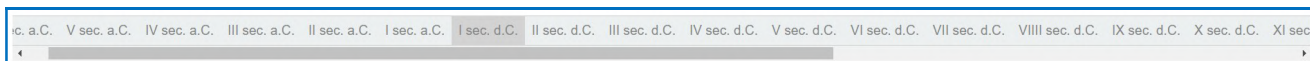


Fig. 3 - Timeline per la ricerca delle epigrafi per secoli.

La ricerca sui testi delle iscrizioni può essere effettuata secondo due diverse modalità:

- la *ricerca libera*, in cui è possibile ricercare un'arbitraria stringa di testo;
- la *ricerca con dizionari*, in cui le parole da ricercare possono essere selezionate da *liste-dizionario* built-in precompilate. Tra le liste-dizionario troviamo, in particolare, le liste dei *nomi gentilizi* e dei *cognomina* sia delle iscrizioni catanesi, sia delle urbane (si veda [2]).

Sono supportate inoltre ricerche di tipo case-sensitive, ricerche per sottostringhe, nonché ricerche per esclusione attraverso l'uso dell'operatore di negazione NOT. Si noti, altresì, che le ricerche libere sui testi delle iscrizioni sono facilitate dall'utilizzo, facoltativo, di una tastiera virtuale (creata appositamente per il progetto) che permette di inserire i vari simboli dell'alfabeto greco con relativi segni diacritici (fig. 4).



Fig. 4 - Tastiera virtuale per l'inserimento dei caratteri greci.

Le ricerche sugli altri campi possono essere effettuate selezionando i termini da apposite liste in forma di menù a tendina. Nel caso dei campi materiale, tipo di oggetto, tipo di iscrizione, tecnica di esecuzione, lingua dell'iscrizione, tali liste formano sottoinsiemi dei vocabolari controllati definiti nel progetto EAGLE (Electronic Archive of Greek and Latin Epigraphy)<sup>6</sup>.

Ogni operazione di ricerca produce come risultato una lista di epigrafi visualizzata in forma di tabella HTML. Si può poi passare alla scheda completa in HTML di ogni singola epigrafe con tutte le informazioni che la descrivono, con il testo e le foto e con la possibilità di visualizzare e scaricare le codifiche in EpiDoc XML. Si osservi che il testo delle iscrizioni viene visualizzato riproducendo la simbologia standard adottata dalle convenzioni Leiden.

Sul lato *back-end*, il *core* del sistema di ricerca si basa sull'indicizzazione dei campi rilevanti di ricerca, ottenuta attraverso uno specifico *mapping* rappresentato da file JSON

<sup>6</sup> Accessibili all'indirizzo web <https://www.eagle-network.eu/resources/vocabularies/>.

contenenti coppie del tipo (fieldName, fileList), dove fileList è la sequenza (opportunamente) ordinata dei nomi dei file XML “corrispondenti” al campo fieldName. Si osservi che l'ordinamento dei nomi di file XML risulta particolarmente utile per accelerare l'esecuzione delle *query* di ricerca. Infatti, mediante l'utilizzo di appositi algoritmi che sfruttano tale ordinamento, i *result-set* delle query possono essere computati in tempo lineare.

**PERCORSI** permette di navigare attraverso le epigrafi secondo i seguenti percorsi tematici: Collezione di provenienza, Lingua, Luogo di provenienza, Ambito religioso, Periodo, Copie, Onomastica, Bibliografia. Ad esempio in Ambito religioso si possono visualizzare tutte le epigrafi cristiane, pagane, ebraiche oppure non identificabili. Mentre Bibliografia consente una ricerca delle iscrizioni a partire dalla bibliografia presente nella scheda epigrafica.

**PROGETTO** con le sottovoci DESCRIZIONE, PARTNER, DOCUMENTAZIONE, MOSTRA e CONVEGNO.

- Nelle pagine **DESCRIZIONE**, **PARTNER** e **DOCUMENTAZIONE** viene presentato il progetto EpiCUM con la descrizione degli enti coinvolti e le pubblicazioni e la rassegna stampa che ha trattato il progetto e la mostra *Voci di Pietra*.
- **MOSTRA** descrive appunto i contenuti della mostra *Voci di Pietra* fornendo anche l'accesso alle pagine web del *chiosco multimediale* appositamente creato per la mostra al fine di favorirne l'approfondimento della visita attraverso la navigazione tra le informazioni raccolte nelle schede, nei video e nelle immagini digitalizzate, sia in italiano che in inglese.
- **CONVEGNO** visualizza una pagina con il programma del convegno scientifico internazionale “*Voci di Pietra*”. *Pluralismo culturale e integrazione nella Sicilia antica e tardo-antica*, organizzato all'interno delle attività del progetto.

## SVILUPPI FUTURI

Coerentemente con le nuove tendenze di promozione del patrimonio culturale attraverso il digitale e le tecnologie informatiche, è prevista nella sezione open data del museo digitale la rappresentazione semantica secondo il paradigma dei Linked Open Data dei concetti dei testi epigrafici, organizzandola in un'unica ontologia del patrimonio epigrafico del museo civico, corredata da uno SPARQL endpoint per interrogare il dataset. È già stato definito lo schema ontologico delle epigrafi, ma si stanno ancora realizzando gli opportuni mapping per popolarlo con i dati dei file XML.

## RINGRAZIAMENTI

Il progetto EpiCUM è stato finanziato dal Patto per Catania a valere sul Fondo Sviluppo e Coesione 2014-2020: Piano per il Mezzogiorno. Si ringrazia Jonathan Prag, responsabile del progetto I.Sicily, per la preziosa collaborazione e per la supervisione dell'edizione digitale della collezione. Si ringrazia inoltre Gaetano Carmelo La Delfa per il supporto nella realizzazione del front-end del museo digitale.

## BIBLIOGRAFIA

- [1] Agodi, S., Cristofaro, S., Noto, V., Prag, J., Spampinato, D. 2018. *Una collaborazione tra museo, enti di ricerca e scuola: l'epigrafia digitale e l'alternanza scuola-lavoro*. *Umanistica Digitale*, 1(1): 207-224, 2018.

- [2] Korhonen, K. 2004. *Le iscrizioni del Museo Civico di Catania: Storia delle collezioni - Cultura epigrafica - Edizione*. Societas Scientiarum Fennica: Commentationes humanarum litterarum, vol. 121, Helsinki, pp. 418.

# Digital Humanities and “Niche” Research Fields: The Case of Ancient Arabian Epigraphy

Irene Rossi

Consiglio Nazionale delle Ricerche, Italy, irene.rossi@cnr.it

## ABSTRACT

The paper aims to contribute to the discussion of the challenges, impact and prospects of digital initiatives engaged with “young” and “niche” sectors of research, through the experience of DASI – *Digital Archive for the Study of pre-Islamic Arabian Inscriptions*. Formerly an ERC-funded project at the University of Pisa in partnership with Scuola Normale Superiore di Pisa, DASI is now maintained thanks to the cooperation between the University of Pisa and the CNR. The objective of DASI is to foster the study of the enormous epigraphic heritage of pre-Islamic Arabia, which spans from the beginning of the I mill. BCE to VI cent. CE. Attesting different Semitic languages and scripts, epigraphs are the only textual sources for the study of Ancient Arabia. Relevant Digital Humanities theoretical and methodological challenges encountered during the course of the project (e.g. issues of choices in data modelling in relation to the complex nature of the epigraphic source, of harmonization of definitions for “marginal epigraphies”, of applying language technologies to under-resourced languages’ corpora, of using the electronic resources for didactic aims) are discussed, in relation to the description of DASI’s digital tools for archiving, curated edition, online fruition, and cultural and linguistic research. Finally, the research prospects of the project, in terms of widening contents, updating technologies and disseminating results, are considered against issues such as the sustainability of short-term funded projects, and of the evaluation of digital, “non-traditional” outputs and publications in academic careers.

## KEYWORDS

Digital Epigraphy, Ancient Near East, pre-Islamic Arabia, curated digital edition, harmonization, digital lexicography, teaching, sustainability.

## (DIGITAL) EPIGRAPHY OF PRE-ISLAMIC ARABIA

The Arabian peninsula in ancient times was inhabited by different civilizations that produced, from the early 1<sup>st</sup> millennium BCE until the advent of Islam, a written documentation consisting of tens of thousands of inscriptions in different Semitic languages (Ancient South Arabian and Ancient North Arabian languages, Old Arabic, and Aramaic varieties – mainly Nabataean) and alphabetic scripts (varieties of both the South Semitic and the Phoenico-Aramaic script families; see Macdonald 2000) (fig. 1). The history of the studies on Ancient Arabian epigraphy is quite recent, as it began towards the end of the 19<sup>th</sup> century (Avanzini 2016), and has been the interest of few research centres, remaining a niche sector even in Ancient Near Eastern studies and Semitic linguistics. The exclusively epigraphic nature of the textual sources, and the complexity of reconstructing the historical, linguistic and cultural scenario on the basis of a fragmentary evidence, have solicited – and at the same time posed challenges to – the digital approach to such an *in fieri* discipline.

Although the first initiatives of digitization of the pre-Islamic Arabian inscriptions were launched already in the 1990s, it was not until the 2010s that they gained consideration by funding institutions. On the basis of the experience of the project *Corpus of South Arabian inscriptions* - CSAI, recently expanded into the *Digital Archive for the Study of pre-Islamic*

*Arabian Inscriptions* - DASI ([www.dasi.cnr.it](http://www.dasi.cnr.it)) – funded from 2011 to 2016 through an ERC Grant awarded to Prof. A. Avanzini of the University of Pisa, in partnership with Scuola Normale Superiore di Pisa – this paper aims at contributing to the discussion on the challenges, impact and prospects of digital initiatives engaged with “young” and “niche” sectors of the humanities research.



Figure 1 – Ancient South Arabian inscription (Şan ‘ā’, Military Museum, 3630)

## DESIGNING DIGITAL TOOLS FOR UNDER-RESOURCED EPIGRAPHIES: CHALLENGES AND SOLUTIONS

*Modelling* – When defining good practices in the digitization process of the epigraphic documents, the double nature (both textual and material) of the (re)source “inscription” must be acknowledged (Eck 2017). However, the stronger emphasis traditionally put on the textual aspect of an inscription led the majority of epigraphic digital corpora to concentrate annotation efforts on the text of the inscription and to treat the contextual information as accessory information. This kind of approach found in the XML a suitable technology; accordingly, the CSAI, as many other epigraphic projects of the 20<sup>th</sup> century, was conceived as an XML-based edition of inscriptions (Avanzini et al. 2000). However, the study of under-resourced corpora, which is typical of “niche” and *in fieri* disciplines, has especially highlighted the limitations of this approach, as the thorough understanding of an epigraphic source implies the integration of contextual information (*in primis* geographical, archaeological, historic-artistic), which is better reached via a database approach. This was then preferred while designing the new archiving system of the project DASI: its conceptual model is centred on the two entities containing information on metadata and text of the inscription and on its physical carrier respectively, completed by entities for: their translations; their sites of origin or provenance; images; bibliographical references; editorial responsibilities (fig. 2). The capability of the web-based database to manage a distributed workflow also allows a collaborative and specialized approach to the epigraphic work.

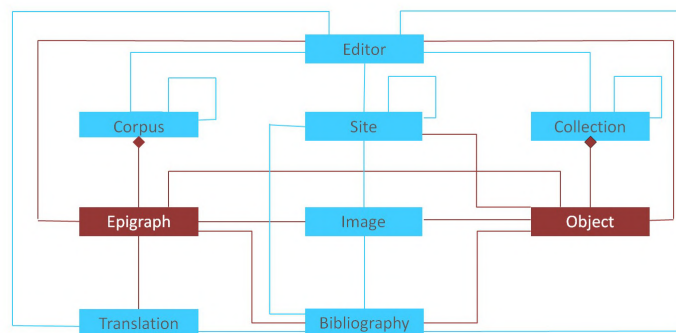


Figure 2 – DASI conceptual model

Indeed, the advantages in flexibility and interoperability allowed by an XML annotation of the text have been preserved by the integration of a module for the Latin transliteration of the text, the encoding of various philological and linguistic phenomena in XML-EpiDoc, and the critical apparatus' annotation (fig. 3).

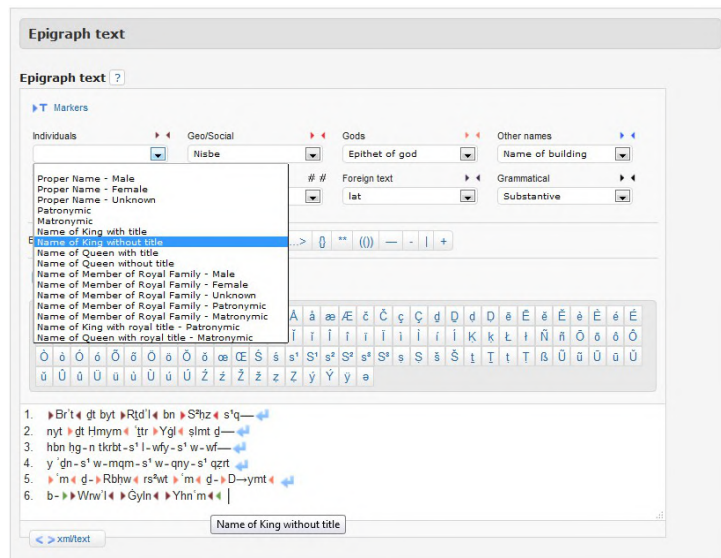


Figure 3 – DASI module for textual annotation

*Interoperability and harmonization* – CSAI had developed its own schema and vocabularies, as many other projects did when standards and thesauri were not yet widespread. At a later stage, in the need for interoperability, marginal epigraphies have found ready-to-use standards and guidelines that had been defined for Classical cultures. This has entailed a considerable effort of harmonization of definitions and phenomena that could only partially be applied to describe sectorial specificities. For instance, adhering to the EpiDoc guidelines (Elliott et al. 2007-2016) was not a straightforward operation. This effort of alignment was a heuristic process and allowed to rethink and refine the CSAI schema, but at the same time highlighted the constraints imposed by legacy data, forcing to compromise solutions, and the need for preserving the cultural peculiarities and the specific tradition of studies (e.g. in the choice and representation of phenomena to annotate) (Avanzini et al. 2016). While technical interoperability is possible (a repository allows service providers to harvest DASI records through the OAI-PMH protocol), reaching interoperability at the content level is more complicated for *in fieri* subjects. This has been apparent in the effort of harmonization of the vocabularies related to metadata – such as the definition of the text genre, or of the typology of support – to thesauri such as the EAGLE Vocabularies (<https://www.eagle-network.eu/resources/vocabularies/>) or the Getty AAT (<https://www.getty.edu/research/tools/vocabularies/aat/index.html>) (Avanzini et al. 2014).

*Lexicography* – Interoperability at a linguistic level across different corpora remains a desideratum, at least until they are extensively “lemmatized”, and translations of single “lemmas” and of entire texts are provided, as well as indications of their semantic categories. To this aim, a tool for creating lexica of the under-resourced languages considered within DASI has been implemented (fig. 4), facing the challenges posed by the ambiguous morphology of the Semitic languages, by the fragmentary nature of the epigraphic sources, and by the current paucity of established linguistic reference tools for the languages of Ancient Arabia (Avanzini et al. 2015). Waiting for the completion of the digital lexica, complex linguistic queries on the texts of the inscriptions can be already performed in the openly accessible DASI front end site, also in combination with metadata, in order to better



appreciate the message of the epigraphic texts against their context. Because of the frequent uncertainties in the reading of the texts, as well as in the interpretation of lexical and onomastic features, search on textual variants has been enabled, thanks to their encoding in the inline apparatus of the inscriptional text.

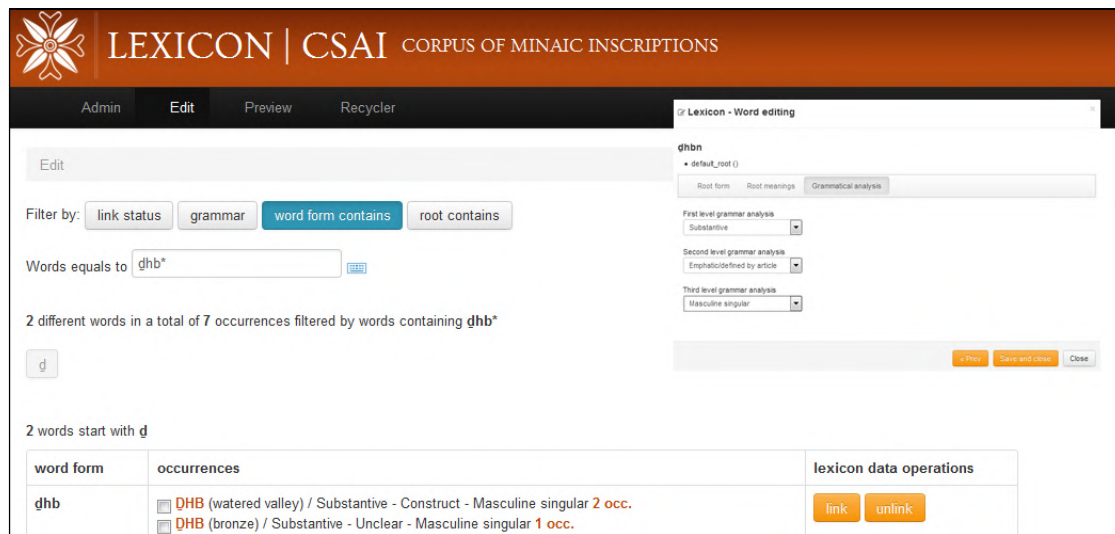


Figure 4 – DASI Lexicon

*Teaching* – Working on ancient Semitic epigraphic languages can be slow and discouraging. In the attempt to get early stage researchers and students closer to the discipline, and to make young generations appreciate the importance of the first-hand approach to ancient primary sources, the validity of DASI system as a didactic tool was tested in the frame of three DASI international training courses and in the teaching of Ancient South Arabian epigraphy at the University of Pisa, with positive results. Stimulated by the interactive method, and thanks to the user-friendly data entry interface, the students were able to easily complete the whole process of curation and edition of an inscription. The pre-defined fields to be compiled guide the analysis of the relevant textual and contextual features; the lists of controlled terms suggest coherent definitions to get used to. The textual encoding gets users familiar with methods and tools of philological analysis and enables linguistic practice through grammatical and onomastic annotations. Eventually, the results of the editorial work are published in the openly accessible online archive, providing immediate and concrete feedback.

Based on our experience, we believe that a more structured programming of teaching the fundamentals of Digital Humanities in the Italian graduate courses would improve the usability, in didactic activities, of the many digital tools specifically designed for research aims, leading to a wider and earlier involvement of students in research projects (on teaching digital epigraphy, see also Bodard and Stoyanova 2016).

## PROSPECTS

DASI project's primary aim was to enhance the study of ancient Arabia through the publication of its curated epigraphic corpora, with a main historical and linguistic concern. The open accessibility of DASI's contents (reaching nearly 8,200 epigraphic and 600 more anepigraphic objects) has contributed to the dissemination of the knowledge on Ancient Arabia beyond the specialists' circuit.

The implementation of DASI archive, both in terms of content and technologies, in the last three years has faced the difficulties that every project – especially if concerned with niche

studies – encounters when its economic support comes to an end. The problem of sustainability is particularly sensitive in Digital Humanities research, to the extent of being one of the reasons of the still poor value attributed to digital products, and specifically to online archives' curated editions, in the academic evaluation process. However, the very recent migration of DASI's IT system to the CNR's infrastructures opens a new phase in the life of the archive, allowing further Digital Epigraphy developments.

In fact, in the meantime, new prospects have opened from the reflection on best practice in the digital approach to the epigraphic sources, thanks to the theoretical and methodological confrontation among Digital Epigraphy projects (e.g. Orlandi et al. 2014, 2017). Some recent editorial initiatives dealing with “marginal” epigraphies (cf. Felle and Rocco 2016) and with the Ancient Near Eastern textual heritage (cf. Bigot Juloux et al. 2018; Cotticelli-Kurras and Giusfredi 2018; De Santis and Rossi 2018) are opening perspectives for a renovation of the discipline, because of the peculiarities of the sources and the different research questions at the basis (with respect to digital Classical epigraphy). Research on methodologies and technologies for cataloguing, fruition and data mining is pushed further by *in fieri* studies, where uncertainty, ambiguity and variation have to be assumed.

As a first step, DASI will explore the application of semantic technologies to its data, starting from the geographic information of its archive contained both in the Site records and in the inscriptional texts. The creation of a gazetteer is envisaged in the frame of the partnership with the recently funded French project MAPARABIA, based at the CNRS, which will use DASI data for the building of a historical atlas and an online encyclopaedia of Ancient Arabia. Integration of other existing digital resources, especially in a Linked Open Data perspective, will be explored in view of the dissemination of the results achieved and the re-use of contents, as well as the application of machine learning techniques to the pre-Islamic Arabian epigraphic material.

## ACKNOWLEDGEMENTS

The research leading to the results presented in this paper has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n° 269774.

## REFERENCES

- [1] Avanzini, A. 2016. *By Land and by Sea: A History of South Arabia before Islam Recounted from Inscriptions*. Roma: «L'Erma» di Bretschneider.
- [2] Avanzini, A., De Santis, A., Gallo, M., Marotta, D. and Rossi, I. 2015. *Computational Lexicography and Digital Epigraphy: Building digital lexica of fragmentary attested languages in the Project DASI*. In: G. Guidi, R. Scopigno, J.C. Torres and H. Graf (eds.), *2015 Digital Heritage International Congress*. New York: IEEE. Pp. 405-408. doi: 10.1109/DigitalHeritage.2015.7419535
- [3] Avanzini, A., De Santis, A., Marotta, D. and Rossi, I. 2014. *Between harmonization and peculiarities of scientific domains: Digitizing the epigraphic heritage of the pre-Islamic Arabia in the project DASI*. In: Orlandi, S., Santucci, R., Casarosa, V. and Liuzzo, P.M. (eds.), *Information Technologies for Epigraphy and Cultural Heritage: Proceedings of the First EAGLE International Conference*. Roma: Sapienza Università Editrice. Pp. 69-93. Retrieved from [<https://www.eagle-network.eu/wp-content/uploads/2015/01/Paris-Conference-Proceedings.pdf>], 2018/11/11.
- [4] Avanzini, A., De Santis, A., Marotta, D. and Rossi, I. 2016. *Is still Arabia at the margins of digital epigraphy? Challenges in the digitization of the pre-Islamic inscriptions in the project DASI*. In A.E. Felle and A. Rocco (eds.), *Off the beaten*

- Track: Epigraphy at the Borders. Proceedings of the VI EAGLE International Meeting (24-25 September 2015, Bari, Italy). Oxford: Archaeopress. Pp. 46-59. Retrieved from [www.archaeopress.com\\ArchaeopressShop\\Public\\download.asp?id={E7B2AAC6-9986-4C41-9842-6AA93BE7ACD9}], 30/11/2017.
- [5] Avanzini, A., Lombardini, D. and Mazzini, G. 2000. *Corpus of South Arabian Inscriptions. La pubblicazione integrale del corpus sudarabico qatabanico*. In: Bollettino d'Informazioni del Centro Informatico per i Beni Culturali della Scuola Normale Superiore 10.
- [6] Bigot Juloux, V., Gansell, A.R. and Di Ludovico, A. (eds.) 2018. *CyberResearch on the Ancient Near East and Neighboring Regions. Case Studies on Archaeological Data, Objects, Texts, and Digital Archiving*. Leiden: Brill. Retrieved from [https://brill.com/downloadpdf/book/edcoll/9789004375086/B9789004375086\_001.xml], 2018/11/11.
- [7] Bodard, G. and Stoyanova, S. 2016. Epigraphers and Encoders: Strategies for Teaching and Learning Digital Epigraphy. In: G. Bodard and M. Romanello (eds.), *Digital Classics Outside the Echo-Chamber: Teaching, Knowledge Exchange & Public Engagement*. London: Ubiquity Press. Pp. 51-68. doi: 10.5334/bat.d
- [8] Cotticelli-Kurras, P. and Giusfredi, F. (eds.) 2018. *Formal Representation and the Digital Humanities*. Cambridge: Cambridge Scholars Publishing.
- [9] De Santis, A. and Rossi, I. (eds.) 2018, in press. *Crossing Experiences In Digital Epigraphy: From Practice To Discipline*. Warsaw-Berlin: De Gruyter.
- [10] Eck, W. 2017. Tradition and progress. The Roman World in the Digital Age - seen through Inscriptions. In: S. Orlandi, R. Santucci, F. Mambrini and P.M. Liuzzo (eds.), *Digital and Traditional Epigraphy in Context. Proceedings of the EAGLE 2016 International Conference*. Roma: Sapienza Università Editrice. Pp. 19-36. Retrieved from [http://www.editricesapienza.it/sites/default/files/5533\_EAGLE\_2016\_International\_Conference.pdf], 2018/11/11.
- [11] Elliott, T., Bodard, G., Mylonas, E., Stoyanova, S., Tupman, C., Vanderbilt, S. et al. 2007-2016. *EpiDoc Guidelines: Ancient documents in TEI XML (Version 8)*. Retrieved from [http://www.stoa.org/epidoc/gl/latest/], 2018/11/11.
- [12] Felle, A.E. and Rocco, A. (eds.) 2016, *Off the beaten Track: Epigraphy at the Borders. Proceedings of the VI EAGLE International Meeting (24-25 September 2015, Bari, Italy)*. Oxford: Archaeopress. Retrieved from [www.archaeopress.com\\ArchaeopressShop\\Public\\download.asp?id={E7B2AAC6-9986-4C41-9842-6AA93BE7ACD9}], 2018/11/11.
- [13] Macdonald, M.C.A. 2000. *Reflections on the linguistic map of pre-Islamic Arabia*. In: *Arabian Archaeology and Epigraphy* 11, 28-79.
- [14] Orlandi, S., Santucci, R., Casarosa, V. and Liuzzo, P.M. (eds.) 2014, *Information Technologies for Epigraphy and Cultural Heritage: Proceedings of the First EAGLE International Conference*. Roma: Sapienza Università Editrice. Retrieved from [https://www.eagle-network.eu/wp-content/uploads/2015/01/Paris-Conference-Proceedings.pdf], 2018/11/11.
- [15] Orlandi, S., Santucci, R., Mambrini, F. and Liuzzo, P.M. (eds.) 2017, *Digital and Traditional Epigraphy in Context. Proceedings of the EAGLE 2016 International Conference*. Roma: Sapienza Università Editrice. Retrieved from [http://www.editricesapienza.it/sites/default/files/5533\_EAGLE\_2016\_International\_Conference.pdf], 2018/11/11.

# Gli archivi personali nell'era digitale: stato dell'arte, criticità e prospettive

Stefano Allegrezza

Università degli studi di Bologna, stefano.allegrezza@unibo.it

## ABSTRACT

L'intervento si propone l'obiettivo di proporre una riflessione critica sulle conseguenze che l'avvento del digitale sta avendo sull'archiviazione e conservazione degli archivi di persona. In un passato neanche troppo remoto i documenti presenti in tali tipologie di archivi si sono sedimentati facendo uso di supporti scrittori (come la carta) che, pur nella loro "fragilità" hanno consentito che i documenti giungessero fino a noi. Spesso si sono verificati ritrovamenti casuali e fortuiti, come nel caso delle fotografie abbandonate per anni nelle soffitte o nei rispostigli, magari risposte in scatole o contenitori di fortuna, e ritrovate – perfettamente leggibili – a distanza di anni dal momento in cui il soggetto produttore le aveva create.

Sfortunatamente, non possiamo dire la stessa cosa per i documenti digitali, dal momento che essi sono costituiti, in ultima analisi, da sequenze di bit che sono ben più difficili da conservare dei loro corrispettivi cartacei, dovendo contrastare il fenomeno dell'obsolescenza che coinvolge le piattaforme tecnologiche necessarie per "leggere" tali documenti, i supporti di memorizzazione su cui sono archiviati ed i formati elettronici secondo i quali vengono codificati. Tutto questo non può non porci dinanzi ad un interrogativo: cosa rimarrà degli archivi di persona che si stanno formando oggi? Quali memorie lasceremo ai nostri posteri?

Se si considera l'importanza dell'immenso patrimonio digitale che si sta formando e che costituirà una fonte privilegiata di informazioni per tutta una serie di settori scientifici (nella ricerca storica, demo-etno-antropologica, sociologica, politica, filologica, artistica, geografica, letteraria, solo per citarne alcuni) si comprende come l'eventualità della perdita degli archivi personali digitali che si stanno formando oggi avrebbe conseguenze estremamente negative per molteplici settori della ricerca. Su questi temi è urgente cominciare ad acquisire consapevolezza per poi arrivare a suggerire soluzioni concrete in grado di assicurare la conservazione nel tempo degli archivi personali digitali.

## KEYWORDS

archivi, archivi digitali, archivi di persona, archivi personali, personal digital archiving

## INTRODUZIONE

Negli ultimi venti anni la "rivoluzione digitale" ha completamente trasformato molte delle attività che per secoli si sono svolte facendo uso di un supporto scritto cartaceo. o, comunque, "analogico", trasformando, di conseguenza, anche gli archivi personali. Si considerino, ad esempio, le tradizionali lettere cartacee, quelle che in un passato neanche troppo remoto venivano scritte a mano, con la macchina da scrivere o con il computer (in quest'ultimo caso occorre poi stamparle), affrancate ed infine imbucate: negli ultimi dieci-venti anni sono state letteralmente "soppiantate" dalle e-mail che risiedono sui sistemi di posta elettronica e si leggono con gli smartphone (e, quindi, non necessitano più di essere stampate). Si pensi alle cartoline che sono state sostituite prima dalle cartoline virtuali che si "spedivano" sul web e poi dai selfie scattati nelle località turistiche più disparate e pubblicati sui social media (ad esempio: Facebook) o su WhatsApp. Si pensi al taccuino su cui nel

passato si prendevano appunti e note oggi sempre più spesso viene sostituito da una delle tante “App” che consentono di prendere appunti elettronici direttamente sugli smartphone. Ancora: i telegrammi che nel passato venivano utilizzati per comunicare notizie in tempi rapidi (si pensi ai telegrammi di condoglianze o di felicitazioni) oggi sono stati “rimpiazzati” da altri strumenti che consentono di inviare comunicazioni pressoché in tempo reale (si pensi, ad esempio, agli SMS) e di avere quasi istantaneamente la garanzia della consegna al destinatario (si pensi ai messaggi inviati tramite WhatsApp); la “vecchia” agenda cartacea è stata soppiantata dalle varie agende elettroniche oggi sono disponibili, a partire da quelle presenti sul proprio telefono cellulare o smartphone per arrivare a quelle accessibili sul web (si pensi a Google Calendar o a Microsoft Office 365, tanto per citare alcune tra quelle più utilizzate) ed accessibili attraverso i propri dispositivi informatici (personal computer, notebook, tablet o smartphone). Allo stesso modo il “vecchio” diario che nel passato raccoglieva, quasi un amico silenzioso, i pensieri più intimi, è stato oggi soppiantato dai diari virtuali – il caso più famoso è quello di Facebook – che, all’opposto, è diventato il modo per rendere noto a tutti la propria vita personale in una sorta di esasperato protagonismo individuale. Persino gli appunti presi durante le riunioni di lavoro o le lezioni oggi sempre più spesso vengono scritti direttamente con il computer, con il tablet o, addirittura, con lo smartphone e, quindi, sono in formato digitale.

Nel mondo multimediale la situazione non è differente: dall’inizio del secolo le fotografie analogiche sono state “soppiantate” da quelle digitali, e ormai si stampa su carta solo una piccola frazione delle fotografie che, a centinaia se non a migliaia, vengono scattate: ormai la parola d’ordine è “condivisione” e le foto vengono immediatamente pubblicate sui social media che, come Facebook, Instagram e Flickr, hanno conosciuto un enorme sviluppo proprio in questi ultimi anni. Allo stesso modo, le registrazioni audio e video che “catturano” i momenti importanti della nostra esistenza sono ormai esclusivamente in formato digitale.

Si tratta solo di alcuni degli innumerevoli esempi<sup>1</sup> che si potrebbero fare, ma sono sufficienti per comprendere quale sia la portata della rivoluzione digitale che stiamo vivendo e alla quale, nel bene o nel male, ci stiamo abituando ed adeguando. Ed è chiaro il trend evolutivo verso il quale anche gli archivi di persona si stanno dirigendo: è molto probabile (se non quasi certo) che in un futuro neanche troppo lontano questi archivi – e non solamente questi – diventeranno totalmente digitali. Inoltre, entreranno a far parte di questa particolare categoria di archivi nuove tipologie di “documenti” che nel passato non esistevano, come i siti web personali, i *blog*, i profili su Facebook, gli *account* su Twitter, Instagram, Pinterest, etc. Se qualcosa rimarrà di analogico, saranno i documenti che, formati in digitale (e, quindi, nativi digitali), verranno occasionalmente stampati per motivi contingenti e cestinati non appena si esaurisce la necessità della loro conservazione. È in atto una vera e propria rivoluzione paragonabile secondo alcuni a quella conseguente all’invenzione della stampa<sup>2</sup>.

Tutto questo ha delle conseguenze - del tutto inimmaginabili fino a pochi anni fa - sulla conservazione a lungo termine degli archivi personali digitali e sulle quali la comunità

---

<sup>1</sup> Viene da chiedersi quali tipologie documentali abbiano, ad oggi, resistito a questa trasformazione digitale. Ad una attenta analisi sono ben poche: alcuni continuano ad utilizzare il diario cartaceo (che forse consente di ricostruire le attività fatte nel corso della propria vita meglio del diario on-line – si pensi a Google Calendar – o sul proprio telefono cellulare); altri continuano ad utilizzare block-notes o quaderni per prendere appunti durante lezioni e corsi; infine una tipologia che resiste al processo di digitalizzazione è la cosiddetta “pergamena” che attesta la fine di un percorso di studi (come la pergamena con il voto di laurea) e che viene incorniciata ed appesa alla parete della propria abitazione o dello studio: in quanti rinuncerebbero alla versione cartacea a favore della versione digitale?

<sup>2</sup> Appare evidente che questa “rivoluzione” è in qualche modo legata all’enorme diffusione di dispositivi mobili (smartphone, tablet, phablet, etc.) che hanno fatto sì che sia diventato estremamente semplice per chiunque produrre documenti di varia natura (non solo documenti testuali ma anche immagini, registrazioni audio e video, etc.).

scientifico ha già cominciato a interrogarsi: cosa rimarrà degli archivi di persona che si stanno formando oggi? Quale futuro li attende? E soprattutto: quali memorie lasceremo ai posteri?

## IL PROBLEMA DELLA CONSERVAZIONE DEGLI ARCHIVI DIGITALI

Il problema della conservazione degli archivi digitali – e, nel caso specifico, degli archivi digitali di persona – rientra indubbiamente tra quelli che maggiormente assilla la comunità scientifica degli archivisti. Fino ad oggi la salvaguardia di questi archivi è stata possibile grazie all'adozione di strategie di conservazione e all'utilizzo dei materiali scrittori che, pur nella loro fragilità ed incertezza, hanno consentito che i documenti giungessero fino a noi. Spesso si sono verificati ritrovamenti casuali e fortuiti, come nel caso delle lettere o delle fotografie abbandonate per anni nelle soffitte o nei ripostigli, magari risposte in scatole o contenitori di fortuna, e ritrovate – perfettamente leggibili – a distanza di anni dal momento in cui il soggetto produttore le aveva create. Sfortunatamente, non possiamo dire la stessa cosa per i documenti digitali, dal momento che essi sono costituiti, in ultima analisi, da sequenze di bit che sono ben più difficili da conservare dei loro corrispettivi cartacei, per le ciò per via di almeno tre ordini di difficoltà. In primo luogo i documenti digitali necessitano di risiedere fisicamente su un 'supporto di memorizzazione': può trattarsi del disco fisso del computer dell'utente o quello del *server* del *provider* del servizio di posta elettronica nel caso dei messaggi di posta elettronica; di un disco ottico (come i *compact disc* o i DVD su cui spesso vengono archiviati tali documenti); di una memoria flash (come le *pendrive* che portiamo sempre con noi); di un nastro magnetico (come quelli utilizzati nei *backup* che gli internet *provider* di posta effettuano con regolarità), oppure di uno spazio virtuale come il *cloud*, la 'nuvola' di cui oggi si parla tanto e sulla cui affidabilità stanno indagando progetti di ricerca internazionali come I-Trust [1]. In secondo luogo, i documenti digitali, essendo costituiti in ultima analisi da una sequenza di bit, per poter essere intelligibili devono essere codificati secondo un certo 'formato elettronico'. Infine, mentre per poter leggere, anche a distanza di parecchi anni, un documento cartaceo è sufficiente solamente una buona vista, per poter leggere un documento digitale occorre anche una opportuna 'piattaforma tecnologica' (tipicamente un computer con un certo sistema operativo ed un certo programma di posta, oppure un *tablet* o uno *smartphone*, come avviene oggi sempre più spesso). Purtroppo, queste tre componenti vanno incontro ben presto al problema dell'obsolescenza che rende molto difficile la conservazione nel tempo dei documenti digitali.

## LO STATO DELL'ARTE

A fronte di queste criticità – che richiederebbero un'azione immediata – va rilevato come il dibattito sul *Personal Digital Archiving* si sia sviluppato solo negli ultimi anni, all'incirca dalla metà degli anni 2000, e ha visto negli Stati Uniti la nazione d'origine. La *Library of Congress* è stata ed è tuttora molto attiva negli studi in materia e nelle iniziative di sensibilizzazione degli utenti [3]. È capofila del *National Digital Information Infrastructure and Preservation Program (NDIIPP)* dedicato al tema della Digital Preservation e rende disponibile all'interno del suo sito web una interessante sezione relativa al *Personal Archiving*<sup>3</sup>, attiva all'incirca dal 2007, dove sono liberamente disponibili materiali divulgativi

---

<sup>3</sup> Si veda: <[www.digitalpreservation.gov/personalarchiving](http://www.digitalpreservation.gov/personalarchiving)>. Si veda anche la sezione del sito della *Columbia University* che mette a disposizione interessanti risorse sulle questioni del *Personal Digital Archiving* (ad esempio: informazioni sulla gestione dei supporti di memorizzazione, sulle procedure di *backup* e *recovery*, sulla corretta denominazione dei file, sulle procedure di migrazione, etc.), disponibile all'indirizzo <[library.columbia.edu/locations/dhc/personal-digital-archiving/online-resources.html](http://library.columbia.edu/locations/dhc/personal-digital-archiving/online-resources.html)>; e la sezione del sito dell'*Association for Library Collections and Technical Services (ALCTS)* - una divisione dell'*American Library Association (ALA)* - che mette a disposizione una serie di materiali dedicati al *Personal Digital*

utili per sensibilizzare sul tema della conservazione degli archivi di persona[10]. Lo stesso sito ospita, dal 2011, il blog “*The Signal - Digital Preservation*”, che comprende una categoria dedicata al Personal Archiving. Dal 2010 inoltre la *Library of Congress* contribuisce, insieme ad altre istituzioni e università, all’organizzazione del *Personal Digital Archiving Day*, evento per il quale ha anche reso disponibile online un *Personal Digital Archiving kit*, rivolto a piccole organizzazioni e biblioteche, ma anche ai privati, contenente suggerimenti operativi e risorse per realizzare la manifestazione [6].

La particolarità dell’approccio americano al tema consiste nel taglio concreto e pratico delle iniziative, ma spesso appare eccessivamente basilare ed insufficiente a fornire le informazioni necessarie per una corretta gestione dei propri archivi. Ad esempio, la *Library of Congress* ha emanato delle linee guida in materia di *Personal Digital Archiving* che utilizzano un linguaggio semplice e comune, adatto ad un pubblico eterogeneo e spesso inesperto, che puntano a fornire al singolo individuo informazioni elementari sulla conservazione di foto, audio, video, e-mail, “personal digital records” (ovvero appunti scolastici, ricevute di pagamenti, estratti conto, slide, bozze, etc.) nonché siti web, blog e social media, ma che si rivelano forse troppo semplicistiche e poco approfondite, perdendo in questo modo gran parte della loro utilità. Anche la *Cornell University Library* ha una sezione del proprio sito dedicata al tema del Personal Archiving con una guida sulle strategie e sulle risorse utili nella gestione dei propri archivi personali, nonché una sezione dedicata ai casi reali di perdita di archivi personali denominata molto significativamente “Real Life Horror Stories” [8].

Le *MIT Libraries* mettono a disposizione una serie di informazioni, sotto forma di “opuscoli” e pieghevoli, che contengono utili suggerimenti per la gestione dei propri archivi digitali personali[7].

Di estremo interesse è l’annuale conferenza internazionale sul Personal Digital Archiving, la cui ultima edizione si è tenuta a Houston (Texas) dal 23 al 25 aprile 2018 [5]; essa è l’unica conferenza internazionale specificatamente incentrata sugli archivi digitali personali ed è quindi l’unico momento in cui gli studiosi che si occupano di questi temi possono confrontarsi.

Per quanto riguarda l’Europa va segnalato il progetto PArADIGM (*Personal Archives Accessible in Digital Media*) [4]: condotto dalle Università di Oxford e Manchester negli anni 2005-2007, ha affrontato le problematiche relative alla conservazione di documenti digitali privati attraverso lo studio e la sperimentazione sugli archivi di alcuni politici britannici. Anche il progetto InterPARES3 (2007-2012) [1] si è occupato della conservazione di documenti personali in alcuni casi di studio, focalizzati sull’acquisizione di archivi e collezioni personali da parte di istituzioni archivistiche.

Vale la pena osservare come si tratti, per lo più, di studi che hanno affrontato la questione dell’archiviazione e conservazione degli archivi di persona dal punto di vista dei soggetti che sono istituzionalmente preposti alla custodia della memoria, mentre sono pochi e non esaustivi gli studi che hanno affrontato l’argomento dal punto di vista della persona “comune” (che è il soggetto produttore di tali archivi). Inoltre, eccezion fatta per le linee guida, minimali, emanate dalla *Library of Congress* e per i risultati del progetto PArDIGM, che tuttavia si è limitato ad affrontare prevalentemente una particolare categoria di archivi di persona (quelli degli uomini politici) non vi sono ancora delle linee guida che possano guidare il soggetto produttore (lo studioso, il letterato, il personaggio illustre, il professionista ma

---

*Archiving* al fine di «increase your understanding of common digital files – digital photos, recordings, video, documents, and others – and learn what it takes to preserve them. Technology changes rapidly. If you don’t actively care for your digital possessions you may lose access to them as some technologies become obsolete. Learn about the nature of the problem and hear about some simple, practical tips and tools to help you preserve your digital stuff». I materiali sono disponibili all’indirizzo <[www.ala.org/alcts/confevents/upcoming/webinar/pres/042413](http://www.ala.org/alcts/confevents/upcoming/webinar/pres/042413)>.

anche la persona comune) nelle corrette operazioni di formazione, gestione e conservazione del proprio archivio digitale. Infine, le iniziative attualmente presenti nel panorama internazionale hanno posto l'attenzione sulle questioni – già queste complesse – dell'archiviazione ma hanno trascurato del tutto quelle – ancor più complesse ed urgenti – della conservazione, così come quelle relative alla tutela della privacy e alla riservatezza dei dati.

In conclusione, se da una parte si registra un interesse verso le questioni del *Personal Digital Archiving* che cresce in maniera molto rapida, dall'altra si deve ammettere che il dibattito su questi temi stenta a decollare, manca una riflessione critica sui molteplici aspetti problematici della questione, mancano proposte serie e condivisibili, con la conseguenza che il rischio di una perdita di una parte notevole degli archivi di persona digitali che si stanno formando in questi anni si fa sempre più concreto.

## LE CRITICITÀ DA AFFRONTARE

La questione dell'archiviazione e conservazione degli archivi personali digitali condivide tutte le complesse problematiche relative agli archivi digitali in generale, ma in più presenta alcuni aspetti peculiari che offrono interessanti spunti di studio e di approfondimento:

- la tendenza alla produzione di quantità spesso incontrollabili di documenti, riconducibile da una parte alla disponibilità di strumenti informatici (computer, tablet, ma anche e soprattutto *smartphone*); dall'altra dall'enorme capacità di memorizzazione dei supporti oggi disponibili (si pensi alle *memory card*, ai dischi fissi esterni, ai sistemi basati sul cloud,...) e contemporaneamente dalla diminuzione del costo di tali supporti;
- la mancanza di un criterio logico di ordinamento, dovuto da una parte alla carenza di conoscenze archivistiche (che stanno diventando sempre più importanti nel “nuovo” mondo digitale); dall'altra alla fiducia riposta negli strumenti tecnologici che fanno illudere che sia possibile recuperare quanto necessario semplicemente affidandosi alle funzioni di ricerca le quali, per quanto sofisticate esse siano, non sono comunque in grado di restituire i risultati accurati che si potrebbero ottenere con un archivio correttamente formato; infine, dall'indisponibilità di strumenti archivistici che possano aiutare i soggetti produttori a formare correttamente i propri archivi personali; tutto ciò è di fondamentale importanza dal momento che l'assenza di un criterio di ordinamento conduce alla formazione di un archivio disordinato ed estremamente difficile da “riordinare” in una fase successiva a quella della formazione;
- la tendenza alla dispersione, ovvero a distribuire i documenti digitali su più sistemi di memorizzazione, su più piattaforme, su più servizi; abitudine questa giustificata il più delle volte dalla paura di perdere i documenti (per cui si fanno molteplici copie e gli stessi documenti si trovano duplicati, triplicati, quadruplicati su più supporti) ma che genera inevitabilmente disordine e confusione;
- la mancanza di operazioni di selezione e scarto, attività che invece acquistano un peso crescente a fronte dell'enorme quantità di documenti digitali personali che vengono prodotti e che rischia di diventare velocemente ingestibile;
- la mancanza di consapevolezza sul problema della conservazione digitale e sul fatto che gli archivi digitali di persona sono intrinsecamente più a rischio dei corrispondenti analogici e quindi avrebbero bisogno di una “cura” maggiore.

Non ci sono solo difficoltà di ordine tecnologico: si pensi, ad esempio, al caso della conservazione dei messaggi di posta elettronica presenti, ad esempio, sulla propria webmail: l'accesso alla propria casella di posta elettronica è normalmente protetto da credenziali di



autenticazione (username e password), note solo a colui che ha creato ed utilizza quella particolare casella di posta elettronica. Ma cosa succede alla sua morte? Solitamente risulta pressoché impossibile per chiunque altro recuperare l'accesso alla casella di posta, dal momento che di solito nessuno si preoccupa di comunicare tali credenziali ad altri soggetti: di conseguenza, tutte queste email rimangono "intrappolate" nella casella di posta e ivi rimangono fino a quando, magari dopo un tempo più o meno lungo, l'account viene disabilitato o eliminato dal provider del servizio di posta elettronica "per inattività". In generale, risulta molto difficile (anzi, il più delle volte impossibile) accedere all'archivio digitale di una persona defunta, sia esso costituito dall'archivio delle sue e-mail o da quello dei suoi post su Facebook. Le norme che riguardano la successione dei beni digitali sono sostanzialmente inesistenti e in più di un'occasione i familiari di un defunto non sono riusciti ad accedere al suo account Facebook, ostacolati dall'applicazione pedissequa della legge sulla privacy. In questo modo, sempre più spesso i dati personali, i ricordi e le testimonianze della nostra vita restano conservate esclusivamente in forma di bit e spesso rimangono per sempre celate dietro a password imperscrutabili e protette a volte da legislazioni che non hanno ancora contemplato il problema dell'eredità digitale.

Come si può intuire si tratta di elementi di forte criticità che influiscono fortemente sulle possibilità di successo di una strategia di conservazione applicata agli archivi personali digitali.

## PROSPETTIVE FUTURE

Gli archivi persona costituiscono una categoria di archivi tradizionalmente più a rischio rispetto a quelli appartenenti ad enti ed "organizzazioni" più o meno complesse che, per tutta una serie di motivi (interessi di studio, prescrizioni legislative, rilevanza economica, etc.), ricevono tradizionalmente l'attenzione degli archivisti [11]. Nella loro trasformazione digitale, poi, le criticità sono certamente aumentate. È, quindi, necessario affrontare con urgenza la questione, partendo da uno studio serio e sistematico delle questioni di particolare complessità teorica e tecnica legate alle attività di formazione, gestione e conservazione degli archivi digitali di persona, mettendo in campo interventi di sensibilizzazione per aumentare il livello di consapevolezza su queste tematiche; individuando metodi, sistemi e strategie che siano nel contempo semplici ed efficaci e che siano utilizzabili non solo dagli specialisti del settore ma anche dalle persone comuni.

## BIBLIOGRAFIA

- [1] InterPARES Project, <[www.interpares.org/ip3/ip3\\_index.cfm](http://www.interpares.org/ip3/ip3_index.cfm)>
- [2] InterPARES TRUST Project, <<https://interparestrust.org>>
- [3] Library of Congress (2013), *Perspectives on Personal Digital Archiving*, National Digital Information Infrastructure and Preservation Program, <[www.digitalpreservation.gov/documents/ebookpdf\\_march18.pdf?loclr=blogsig](http://www.digitalpreservation.gov/documents/ebookpdf_march18.pdf?loclr=blogsig)>
- [4] PARADIGM Project <[www.paradigm.ac.uk](http://www.paradigm.ac.uk)>
- [5] Personal Digital Archiving Conference, Houston, TX, April 23-25, 2018, <<https://sites.lib.uh.edu/pda18>>
- [6] "Personal digital archiving kit", <[digitalpreservation.gov/personalarchiving/padKit/index.html](http://digitalpreservation.gov/personalarchiving/padKit/index.html)>
- [7] "Personal digital Archiving strategies", pieghevole disponibile all'indirizzo <[libraries.mit.edu/digital-archives/files/2015/10/2015\\_pda\\_handoutdissemination-v3.pdf](http://libraries.mit.edu/digital-archives/files/2015/10/2015_pda_handoutdissemination-v3.pdf)>
- [8] "Real Life Horror Stories", Cornell University Library <[guides.library.cornell.edu/digitalarchiving](http://guides.library.cornell.edu/digitalarchiving)>

- [9] Redwine, Gabriela (2015) *Personal Digital Archiving* - DPC Technology Watch Report 15-01 December 2015, in: DPC Technology Watch Series, <[www.dpconline.org/docs/technology-watch-reports/1460-twr15-01/file](http://www.dpconline.org/docs/technology-watch-reports/1460-twr15-01/file)>.
- [10] Vettore, Simone (2013), *Archivi digitali di persona: è ora di iniziare a parlarne*, in: Il mondo degli Archivi, Sezione Studi, 9 marzo 2013, <[mda2012-16.ilmondodegliarchivi.org/index.php/studi/item/118-archivi-digitali-di-persona-%C3%A8-ora-di-iniziare-a-parlarne](http://mda2012-16.ilmondodegliarchivi.org/index.php/studi/item/118-archivi-digitali-di-persona-%C3%A8-ora-di-iniziare-a-parlarne)>
- [11] Vettore, Simone (2014), *Gestione degli archivi digitali di persona: strategie e problematiche*, in: Il mondo degli Archivi, Sezione Studi, 4 febbraio 2014, <[mda2012-16.ilmondodegliarchivi.org/index.php/primo-piano/item/283-gestione-degli-archivi-digitali-di-persona-strategie-e-problematiche](http://mda2012-16.ilmondodegliarchivi.org/index.php/primo-piano/item/283-gestione-degli-archivi-digitali-di-persona-strategie-e-problematiche)>

# L'utilizzo dei database da parte degli storici: storiografia e dibattito attuale

Tiago Luis Gil

Università di Brasilia / Università Ca'Foscari di Venezia(Visiting Scholar), Italy, tiagoluisgil@gmail.com

## ABSTRACT

L'obiettivo di questo intervento è esplorare le relazioni tra la storia e l'informatica negli ultimi cinquant'anni attraverso la selezione di testi relativi dell'uso delle banche dati. Il focus sarà esclusivamente rivolto verso i testi già pubblicati sul tema, non considerando, quindi, le ricerche e gli studi che ancora non hanno dato luogo ad articoli o altro tipo di divulgazione scientifica. Questa ricerca si propone di ripercorre le trasformazioni avvenute durante questo periodo di tempo, sia tecniche sia concettuali e semantiche, riservando particolare attenzione a come gli storici hanno si sono relazionati e hanno utilizzato gli strumenti informatici. Come campione sono stati presi in esame centinaia di testi tra libri, capitoli e articoli, proveniente dall'Europa, dal Nord America e altri paesi comprendendo cinque lingue differenti. La conclusione vuole mettere in discussione la prospettiva interdisciplinare degli storici e porre l'accento sulla evidente mancanza di una riflessione sull'uso degli strumenti informatici e digitali.

## PAROLE CHIAVE

Storia ed informatica, interdisciplinarietà, banche dati.

## INTRODUZIONE

L'obiettivo di questo testo è studiare le relazioni tra storia ed informatica negli ultimi cinquanta anni, un lasso di tempo nel quale la storia ha sentito la presenza della giovane disciplina che aveva sedotto tutti i campi del sapere con cui era venuta in contatto. Il testo non vuole rappresentare questa interdisciplinarietà come una tendenza o un dato rappresentativo dell'approccio degli storici, ma come un oggetto di studio in divenire.<sup>1</sup>

Il gruppo campione, composto dagli storici che si sono interessati all'informatica al punto di pubblicare dei testi sull'argomento, non era né dominante né egemonico. Studiarli, tuttavia, può essere utile per fornire alcune idee sulla natura del dibattito e, soprattutto, sulle posizioni prese dagli storici davanti all'avvicinamento delle due discipline. Sarà una storia del consumo di un prodotto specifico, l'informatica, nelle ultime decadi del XX secolo e nelle prime del XXI. Il focus sarà rivolto verso i testi già pubblicati sul tema, non considerando, quindi, le ricerche e gli studi che ancora non hanno dato luogo ad articoli o altro tipo di divulgazione scientifica.

## METODOLOGIA E FONTI

La selezione di fonti (articoli, libri, capitoli) è stata fatta tenendo in conto dell'impossibilità di identificare e studiare le migliaia di ricorrenze di temi informatici nelle pubblicazioni storiche, la maggior parte delle quali non documentate o poco divulgate. Così il campione scelto è abbastanza ridotto e lontano dall'essere esaustivo; tuttavia sono state prese in considerazione 768 pubblicazioni tra il 1967, data dei primi rilevamenti, e il 2010, l'ultimo anno analizzato. Di questo totale, una parte minore, composta da 281 testi, è stata letta e classificata in base ad una lista di tematiche frequenti, che ha permesso di dividere la produzione scientifica in nove temi ricorrenti.

La maggior parte della campionatura è derivata da un numero limitato di dossier e volumi collettanei specificatamente dedicati al tema. Il primo di questi è stato *The Historian and the computer* del 1971<sup>2</sup>. Sempre tra i libri scritti da più autori l'altro grande riferimento è stato *Databases in Historical Research*, del 1996.<sup>3</sup> Più frequenti sono stati i numeri monografici pubblicati in riviste. Nel 1971 usciva il primo numero di «Computers and Medieval Data Processing», una pubblicazione di taglio più informativo che di dibattito, che sarebbe cessata nel 1987. Nel 1979, fu la volta di «Le médiéviste et l'ordinateur», stampata fino al 2006.

I parametri utilizzati per la scelta degli articoli è stato dunque quello di prendere in considerazione quei testi che avessero un riferimento diretto all'uso dell'informatica nel titolo oppure che fossero inseriti all'interno di volumi collettivi specificatamente dedicati al tema. Sono state selezionate espressioni come “computer”, “informatica”, “calcolo automatico dei dati”, “databank”, e le loro variazioni come “database” o “analisi dei dati”; e anche le nuove espressioni come “www”, “internet”, “web”, “multimedia”. Molti degli autori considerati hanno iniziato la loro carriera parlando di “database storici” per poi passare, nel corso del tempo, ad utilizzare espressioni come “digital history”, il termine attualmente in uso; questa analisi ha consentito infatti anche di verificare il cambiamento semantico del campo di indagine.

## **ANALISI DEI DATI**

La parola stessa per designare l'informatica non era molto antica, visto che è stata utilizzata per la prima volta in Germania nel 1957, e in Francia, nel 1962.<sup>4</sup> L'informatica si stava ancora consolidando come disciplina autonoma quando iniziò a suscitare l'interesse tra gli storici. La prima occorrenza di questi termini nel campione analizzato risale al maggio del 1965, nel *Colloque de l'École Normale supérieure de Saint-Cloud*, i cui atti furono pubblicati solo tre anni più tardi. Anche il testo *The historian and the computer*, del 1971, fa riferimento ad alcuni studi nordamericani e inglesi di metà anni sessanta. La rivista «Computers and Medieval Data Processing», nel suo primo numero, nel 1971, dava notizia di una serie di studi di storia medievale che utilizzavano moderne tecniche di analisi di dati, il più vecchio dei quali iniziato nel 1959 e l'altro nel 1964. In linea generale questi dati ci confermano che l'uso dell'informatica da parte degli storici inizia all'inizio degli anni sessanta per poi intensificarsi nella seconda metà della decade.<sup>5</sup>

Durante l'anni 1990 l'uso dell'informatica da parte degli storici raggiunse il suo vertice, grazie all'uso dei PC per l'analisi dei dati. Negli anni novanta si assisté invece ad una flessione come dimostra il calo delle pubblicazioni. Tuttavia, con l'avvento di internet e il massiccio ricorso alle pubblicazioni on line iniziarono ad uscire numerosi articoli, soprattutto nel mondo anglosassone, e più specificatamente negli Stati Uniti, su questi nuovi canali virtuali e sulle possibilità che si stavano aprendo agli storici. La rivista «Journal of the Association for History and Computing», grazie all'ampio spazio dato al dibattito sul ruolo di internet, è un perfetto esempio di questo nuovo scenario. Con l'avvento di internet il calo che si era visto negli anni novanta fu superato da una nuova ondata di interesse.

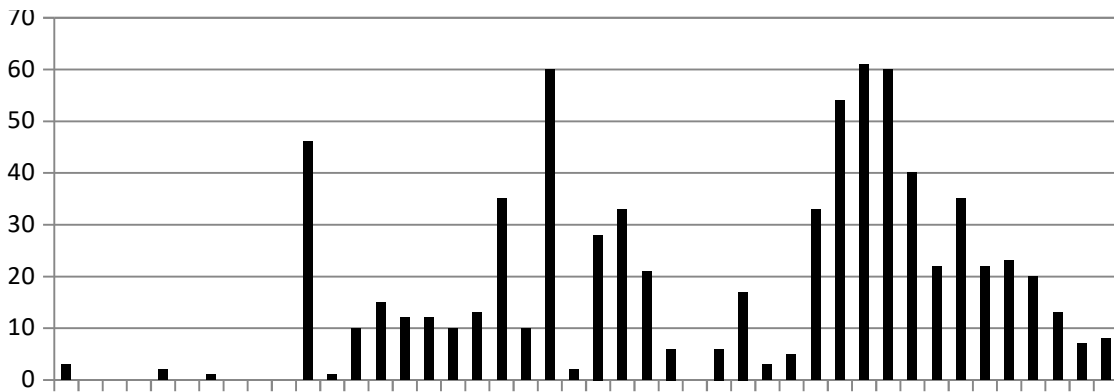


Figura 1 – Fluttuazione annuale delle pubblicazioni che analizzano il rapporto storia/informatica

La sua discontinuità emerge chiaramente in relazione alla natura delle pubblicazioni, alcune delle quali riuniscono decine di contributi che sono stati analizzati singolarmente poiché avevano oggetti di studio differenti tra loro. I dati della figura 1 confermano quanto detto precedentemente: la diffusione di internet ha provocato un forte impatto nel numero delle pubblicazioni e ha permesso la continuità del rapporto tra storia e informatica, anche se in un modo molto differente rispetto ai decenni precedenti. Nel corso dei primi anni del XXI secolo si può notare un'altra inflessione questa volta dovuta ad un nuovo calo dell'utilizzo dell'informatica da parte della storia.

Per fare una analisi più dettagliata, anche se approssimativa, si deve dividere questo lasso di tempo in tre periodi, usando come punti di svolta due avvenimenti che hanno cambiato il consumo informatico: la nascita del *personal computer* alla fine degli anni settanta, e l'avvento di internet a metà anni novanta. A partire da queste premesse, saranno indicati il 1980 e il 1995 come anni di riferimento rappresentanti il culmine di questi due fenomeni. Erano infatti presenti i PC nel periodo precedente alla prima data e internet negli anni antecedenti alla seconda seppure in modo non capillare come dopo i due rispettivi punti di svolta.

È stata fatta anche una ricerca quantitativa sulla ricorrenza di alcune tematiche chiave che sono state al centro del dibattito sul rapporto storia/informatica nel corso del periodo esaminato. Questa ricerca ha aiutato a comprendere meglio quali sono i vari significati che gli storici hanno dato all'informatica. Il campione questa volta è stato più piccolo, riguardando solo un gruppo di 281 pubblicazioni che sono state lette e classificate.

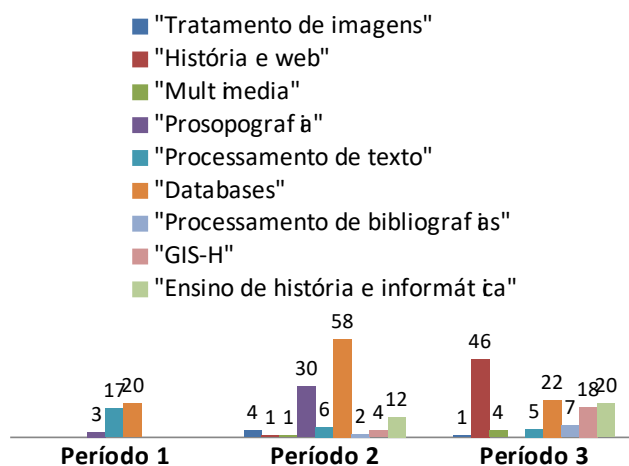


Figura 2 – Variazione delle tematiche legate all’informatica durante i tre periodi (1968-1980; 1981-1995; 1996-2010)

I valori nell’asse X indicano i tre periodi menzionati: 1968-1980, 1981-1995 e 1996-2010. I diversi colori, invece, rappresentano le tematiche. “Trattamento delle immagini” si riferisce agli studi riguardanti l’analisi dell’immagine, ad esempio come in *The automated connoisseur: image analysis and Art history*<sup>6</sup>. “Prosopografia”, è riferito agli studi prosopografici, diffusi in seguito all’articolo di Lawrence Stone del 1977<sup>7</sup>. “Bibliografia” riguarda gli studi bibliometrici come nella recente proposta di Anna Maria Tammaro<sup>8</sup>. Sotto la dicitura “Storia e web” sono stati raccolti i testi con oggetto di studio internet, come il testo di Serge Noiret del 2011<sup>9</sup>. “Analisi del testo” raccoglie tutti gli studi di analisi automatica dei testi, che vanno dalla loro strutturazione all’analisi fattoriale, dalla quantificazione delle parole o alla loro lemmatizzazione. “GIS-H” (*Historical Geographic information system*) indica gli articoli, i libri e i capitoli che trattano del processamento digitale dei dati territoriali storici.<sup>10</sup> Il termine “Multimedia”, che sarebbe potuto essere analizzato con internet, è stato invece analizzato a sé stante. Nella seconda metà degli anni novanta, infatti, queste due tecnologie erano molto diverse e i CD di contenuto multimediale sembravano una tecnologia promettente. “Database” indica l’uso dei database, ed è il settore più longevo e numeroso; esso comprende opere come quelle di Press e Harvey<sup>11</sup>. Infine la categoria “Insegnamento” riguarda quei testi relazionati con l’insegnamento della storia e l’uso di tecnologie informatiche, come l’articolo di MacEachern del 1998.<sup>12</sup>

Il parametro adottato, la ricerca della presenza delle espressioni “storia” e “informatica”, può creare alcuni dubbi sul cambiamento del lessico utilizzato, dato che entrambi i concetti hanno avuto una loro propria evoluzione. Tuttavia sono state riscontrate alcune forti evidenze che hanno indotto a mantenere questo criterio. La più forte di queste è la centrale importanza che hanno avuto gli storici in questo dibattito durante tutto il cinquantennio preso in esame. Uno di questi, il professore dell’Università di Paris I, Jean-Philippe Genet, ha scritto con cadenza regolare sull’argomento a partire dal 1977, ed è anche il coordinatore di una delle ultime opere collettanee prese in considerazione in questo studio edita nel 2011. La sua pluriennale ricerca si è focalizzata sull’analisi di testi, sulla prosopografia, sull’analisi fattoriale e sui database alimentando un continuo scambio con i dibattiti storiografici correnti. Allo stesso modo Manfred Thaller, una figura centrale negli anni ottanta all’interno del coevo dibattito sui database, è presente negli anni duemila in alcuni studi sullo sviluppo di internet. Sono due esempi specifici, è vero, ma entrambi sono stati fondamentali durante il cinquantennio qui preso in esame, come leader di equipe e, a ben guardare, non rappresentano affatto due casi isolati.

## CONSIDERAZIONI GENERALI

Abbiamo analizzato come il dibattito sull'uso dei database nelle studi storici si sia modificato nel periodo preso in considerazione: da una visione ottimista basata sui pregi delle macchine, passando attraverso un atteggiamento molto critico, fino alla perdita di interesse degli ultimi anni. Nel frattempo sono cambiati gli storici ed è cambiata l'informatica; tuttavia, sono riscontrabili delle costanti. L'utilizzo dell'informatica da parte degli storici rimane guidato dalle domande del mercato e dallo sviluppo tecnico, come è avvenuto con la diffusione dei *personal computer* e di internet. I casi in cui è stata la storia a porre delle domande specifiche o a chiedere un adattamento della tecnologia ai fini della ricerca scientifica sono stati lentamente dimenticati, come nel caso del software *kleio* negli anni ottanta, gli studi di Alan Macfarlane negli anni settanta e il dibattito sulle classificazioni iniziato da Daumard negli anni sessanta.<sup>13</sup>

Lo schema presentato sarebbe appropriato per affermare l'interdisciplinarietà degli storici? L'utilizzo dell'informatica da parte della storia può essere comparato con altre contaminazioni disciplinari come quelle che la storia ha avuto con la geografia, l'antropologia e l'economica nel Ventesimo secolo? Saranno solo gli specialisti a testare le metodologie delle altre aree? È possibile che questo atteggiamento poco critico sia da imputarsi, in larga misura, ad una concezione diffusa all'interno della disciplina che sottostima la conoscenza tecnica? Per usare la metafora di Edward Shorter, che paragonava il computer ad una macchina, non sarà più utile sapere prima dove è che si vuole andare con queste macchine?

## BIBLIOGRAFIA

- [1] BRETON, *Une Histoire de L'informatique*, Paris, Seuil, 1990; J.-P. Genet e A. Zorzi, *Les Historiens et L'informatique: Un Métier À Réinventer*, Roma, École Française de Rome, 2011; S. Noiret, *Y a t-il une Histoire Numérique 2.0?*, in *Ibid.*; J.-P. Genet, *Source, Métasource, Texte, Histoire*, in *Storia & Multimedia*, a cura di F. Bocchi e P. Denley, Bologna, 1994.
- [2] SHORTER, *The Historian and the Computer. A Practical Guide*, New Jersey, Prentice-Hall, 1971.
- [3] HARVEY e J. Press, *Databases in Historical Research. Theory, Methods and Applications*, London, Macmillan Press, 1996.
- [4] GIL, Tiago. Our own «in—house» software: una Historia de historiadores programadores. IN: Bresciano y Gil. LA HISTORIOGRAFÍA ANTE EL GIRO DIGITAL Reflexiones teóricas y prácticas metodológicas. Montevideo: Ediciones Cruz del Sur, 2015.
- [5] VAUGHAN, *The Automated Connoisseur: Image Analysis and Art History*, in *History and Computing*, a cura di P. Denley e D. Hopkin, Manchester, Manchester University Press, 1987.
- [6] STONE, *Prosopography*, in «Daedalus», 100, n. 1, 1971, pp. 46–79.
- [7] GREGORY e P.S. ELL, *Historical GIS: Technologies, Methodologies, and Scholarship*, Cambridge Studies in Historical Geography, Cambridge University Press, 2007. [http://books.google.fr/books?id=I\\_HhhJF5Un4C](http://books.google.fr/books?id=I_HhhJF5Un4C).
- [8] HARVEY e J. Press, *Databases in Historical Research*, cit.
- [9] MACEACHERN, *The :-) and :-( of Teaching History on the Web*, in «Journal of the Association for History and Computing», 1, n. 2, novembre 1998.
- [10] GIL, Tiago. Our own «in—house» software: una Historia de historiadores programadores. IN: Bresciano y Gil. LA HISTORIOGRAFÍA ANTE EL GIRO DIGITAL Reflexiones teóricas y prácticas metodológicas. Montevideo: Ediciones Cruz del Sur, 2015.

# SHORT PAPERS



# Analizzare, modellizzare e codificare i pareri di lettura editoriali

Laura Antonietti

Université Grenoble Alpes, Francia – [laura.antonietti@univ-grenoble-alpes.fr](mailto:laura.antonietti@univ-grenoble-alpes.fr)

## ABSTRACT

Questo intervento presenterà i primi risultati dell'analisi, della modellizzazione e della codifica dei pareri di lettura elaborati all'interno della redazione Einaudi nel secondo dopoguerra. Il processo di lettura editoriale dei testi di narrativa, con una particolare attenzione accordata a Natalia Ginzburg, costituisce l'oggetto di studio della mia tesi di dottorato.

Il mio lavoro di ricerca si sviluppa al crocevia tra diverse discipline: la storia del libro e la cultura editoriale, la letteratura italiana contemporanea e le digital humanities. Riguardo a quest'ultimo punto, il mio intervento evidenzierà il contributo fondamentale degli strumenti e del metodo dell'informatica umanistica nell'ambito delle mie ricerche.

Verrà dapprima presentato il modello del processo di lettura einaudiano che ho creato utilizzando il linguaggio UML. Questo mi ha permesso, da una parte, di dare ordine a un corpus disomogeneo, dall'altra ha costituito le basi dell'attività di concettualizzazione necessaria alla realizzazione di un database. Progettato grazie a un adeguato supporto informatico, il database ha permesso di raccogliere in modo strutturato le informazioni per la descrizione del corpus. Questo strumento è stato applicato a un campione di documenti, che sono stati in seguito trascritti e codificati secondo lo schema TEI. È stato elaborato un modello che mette in evidenza le diverse tipologie di documenti, la loro struttura e le relazioni esistenti tra essi. I metadati contenuti nel database e i testi codificati costituiscono le basi del lavoro che seguirà: la creazione di un'edizione digitale che permetterà un'interrogazione trasversale del corpus e ne metterà in luce l'intertestualità.

Se da un lato questo costituisce un fondamentale strumento di riflessione e di analisi, dall'altro l'edizione del corpus ha lo scopo di mettere in luce la specificità e l'importanza che i pareri di lettura hanno nel processo di pubblicazione di un testo e, di conseguenza, nella costruzione del catalogo di una casa editrice.

## PAROLE CHIAVE

Modellizzazione, codifica, UML, XML TEI, lettura editoriale, pareri di lettura.

## ABSTRACT (800-1000 parole)

Questo intervento presenterà i primi risultati dell'analisi, della modellizzazione e della codifica di una selezione dei pareri di lettura elaborati all'interno della redazione della casa editrice Einaudi nel secondo dopoguerra. Questo lavoro rientra nel contesto più ampio del progetto di tesi di dottorato che ho cominciato nell'ottobre 2016 all'Université Grenoble Alpes, in cotutela con l'Università degli Studi di Milano, sotto la guida dei professori Elena Pierazzo e Alberto Cadioli. Il progetto ha come obiettivo quello di analizzare il processo di lettura editoriale che aveva luogo nella casa editrice torinese (Mangoni 1999, Turi 1990), con una particolare attenzione dedicata alla narrativa e al caso di Natalia Ginzburg. Tale lettura veniva condotta all'interno della casa editrice o all'esterno da parte di consulenti per verificare la pubblicabilità di un'opera: quest'attività si concretizzava nella stesura di documenti, che chiamiamo pareri di lettura, nei quali il manoscritto veniva presentato e

valutato. Prodotto di una scrittura professionale che si caratterizza per la sua destinazione e circolazione interna, il parere di lettura rappresenta un passaggio fondamentale del percorso decisionale che porta alla pubblicazione di un'opera e quindi alla costruzione del catalogo di una casa editrice. Attraverso questi documenti i lettori editoriali selezionavano i libri da pubblicare e, di conseguenza, proponevano dei modelli che devono essere interpretati secondo uno specifico orizzonte storico e culturale. Lo studio dei pareri di lettura, in gran parte inediti (una ristretta selezione è stata pubblicata in Munari 2015), costituisce la base della riflessione sull'attività editoriale di Natalia Ginzburg (Iannuzzi 2012; Saita 2009), in particolar modo come lettrice: è stato possibile, d'altro canto, rendersi conto di quanto Natalia Ginzburg, nonostante l'importanza del ruolo assunto presso Einaudi, rimanga a pieno titolo una "protagonista nell'ombra" dell'editoria italiana, per utilizzare una felice espressione coniata da Gian Carlo Ferretti (2012), ossia una di quelle figure che attendono di essere studiate in modo organico e completo per il loro specifico e fondamentale lavoro nell'editoria italiana del Novecento.

Il mio lavoro di ricerca si sviluppa al crocevia tra diverse discipline: la storia del libro e la cultura editoriale, la letteratura italiana contemporanea e le *digital humanities*. Per quel che riguarda in particolare quest'ultimo punto, vorrei mettere in evidenza il contributo fondamentale degli strumenti e del metodo dell'informatica umanistica nell'ambito delle mie ricerche: attraverso questa comunicazione illustrerò come sono arrivata a costruire e a implementare il modello del processo di lettura einaudiano e dei documenti da esso prodotti, sottolineando come la modellizzazione dei dati e la codifica siano da me concepite come attività speculative di grande valore epistemologico (Cummings 2008, Ciula e Eide 2014).

Verrà presentato, in un primo momento, il modello del processo di lettura einaudiano che ho creato, facendo in particolare riferimento ai pareri di lettura individuati principalmente all'interno dell'archivio storico della casa editrice conservato presso l'Archivio di Stato di Torino. Sarà discusso, in particolare, lo schema realizzato proprio allo scopo di descrivere il corpus dei documenti: si tratta precisamente di un diagramma delle classi realizzato con il linguaggio di modellizzazione UML (fig.1). Questo mi ha permesso, da una parte, di dare ordine a un corpus variegato e disomogeneo al suo interno, dall'altra ha costituito le basi dell'attività di concettualizzazione e di analisi necessaria alla realizzazione di un database. Progettato e implementato grazie a un adeguato supporto informatico, il database ha permesso di inventariare le fonti in maniera efficace e performante, raccogliendo in modo strutturato le informazioni fondamentali per la descrizione del corpus. Vorrei segnalare come la realizzazione della base di dati sia da considerarsi non ancora ultimata, in quanto è il frutto di un continuo dialogo tra le esigenze del progetto (che emergono di volta in volta, con il progredire del lavoro di spoglio delle fonti) e le esigenze tecniche di implementazione: riflettere sui metodi e sugli strumenti offerti dall'informatica umanistica è in tal senso uno stimolo per l'analisi e l'interpretazione delle fonti stesse. Il database accoglie allo stato attuale i dati relativi a un campione di documenti, che sono stati in seguito trascritti e codificati secondo lo schema XML TEI. È stato elaborato un modello che mette in evidenza le diverse tipologie testuali (lettere, schede, verbali di riunioni ecc.), la loro struttura e le relazioni esistenti tra essi. I metadati contenuti nel database e i testi codificati costituiscono le basi del lavoro che seguirà: la creazione di un'edizione digitale dei documenti.

La scelta del digitale rimanda a due aspetti diversi: in primo luogo la fruizione dell'edizione. Le edizioni digitali consentono infatti di valorizzare la complessità dei corpora di fonti, mettendo in luce percorsi non immediatamente evidenti ad un'analisi lineare dei documenti, consentendone un'interrogazione trasversale (per parole chiave, tematiche, periodo, ecc.) e valorizzandone l'intertestualità e la multiautorialità, caratteristica insita nei pareri di lettura, testimoni di un processo decisionale a più voci, di un vero e proprio dialogo sulle opere in questione (Einaudi 2001, Munari 2011). L'edizione del corpus ha lo scopo di mettere in

evidenza la specificità e l'importanza che i pareri di lettura hanno nel processo di pubblicazione di un testo e, di conseguenza, nella costruzione del catalogo di una casa editrice, che ha proprio nel catalogo la sua più grande ricchezza e la sua eredità più significativa: non esageriamo quindi se definiamo l'insieme dei pareri di lettura come una parte fondamentale del patrimonio storico dell'editore e la scrittura editoriale come un genere critico particolare e complesso (Bricchi 2010), che merita di essere studiato con appositi strumenti. A tal proposito, vorrei sottolineare come, per quel che riguarda il digitale, le potenzialità dei suoi strumenti devono ancora essere sfruttate e applicate a questo campo di ricerca: allo stato attuale (Franzini, Terras e Mahony 2016, Sahle 2008), non esistono edizioni digitali di pareri di lettura. Tale assenza di modelli diretti di riferimento costituisce una delle sfide più stimolanti del mio progetto. D'altro canto l'utilizzo degli strumenti dell'informatica umanistica rappresenta un aspetto fondamentale del mio lavoro da un punto di vista metodologico: l'attività di modellizzazione è intesa qui come uno strumento conoscitivo, di riflessione e di analisi che costituisce un aspetto chiave dell'attività di interpretazione delle fonti.

#### a. BIBLIOGRAFIA

- [2] Bricchi, M. 2010. «La scheda di lettura come micro-genere letterario». In S. Arduini e I. Carmignani (a cura di), *Giornate della traduzione letteraria. Nuovi contributi*. Roma: Iacobelli, pp. 54-66.
- [3] Ciula, A., Eide, Ø. 2014. «Reflections on cultural heritage and digital humanities: modelling in practice and theory». In First International Conference on Digital Access to Textual Cultural Heritage (DATECH). Madrid, Spain, 2014. Conference Proceedings, pp. 35-41. <<http://modellingdh.uni-koeln.de/wp-content/uploads/2017/01/modelling-inpractice-and-theory.pdf>>.
- [4] Cummings, J. 2008. «The Text Encoding Initiative and the Study of Literature». In R. Siemens et S. Schreibman (a cura di), *A Companion to Digital Literary Studies*. Blackwell. Oxford: Blackwell, pp. 218-239.
- [5] Einaudi, G. 2001. *Tutti i nostri mercoledì*. Bellinzona: Casagrande.
- [6] Ferretti, G. C. (a cura di) 2012. *Protagonisti nell'ombra*. Milano: Unicopli.
- [7] Franzini G., Terras M. e Mahony S. 2016. «A Catalogue of Digital Editions». In M. J. Driscoll e E. Pierazzo (a cura di), *Digital Scholarly Editing Theories and Practices*. Cambridge: Open Book Publishers, pp. 161-182. <[https://www.researchgate.net/publication/264205341\\_A\\_Catalogue\\_of\\_Digital\\_Editions](https://www.researchgate.net/publication/264205341_A_Catalogue_of_Digital_Editions)>.
- [8] Iannuzzi, G. 2012. «Natalia Ginzburg. Una linea editorial-creativa». In G. C. Ferretti (a cura di), *Protagonisti nell'ombra*. Milano: Unicopli, pp. 115-133.
- [9] Mangoni, L. 1999. *Pensare i libri: la casa editrice Einaudi dagli anni Trenta agli anni Sessanta*. Torino: Bollati Boringhieri.
- [10] Munari, T. (a cura di) 2015. *Centolettori: i pareri di lettura dei consulenti Einaudi 1941-1991*. Torino: Einaudi.
- [11] Munari, T. (a cura di) 2011. *I verbali del mercoledì. Riunioni editoriali Einaudi 1943-1952*. Torino: Einaudi.
- [12] Saita, N. 2009. «Natalia Ginzburg la fedeltà di una vita con “passo da soldato”». In R. Cicala e V. La Mendola (a cura di), *Libri e scrittori di via Biancamano, Casi editoriali in 75 anni di Einaudi*. Milano: EDUCatt, pp. 95-124.
- [13] Sahle, P. 2008. *A Catalogue of Digital Scholarly Editions*. <<http://www.digitale-edition.de/>> (ultimo aggiornamento 2018).
- [14] Turi, G. 1990. *Casa Einaudi: libri uomini idee oltre il fascismo*. Bologna: Il mulino.

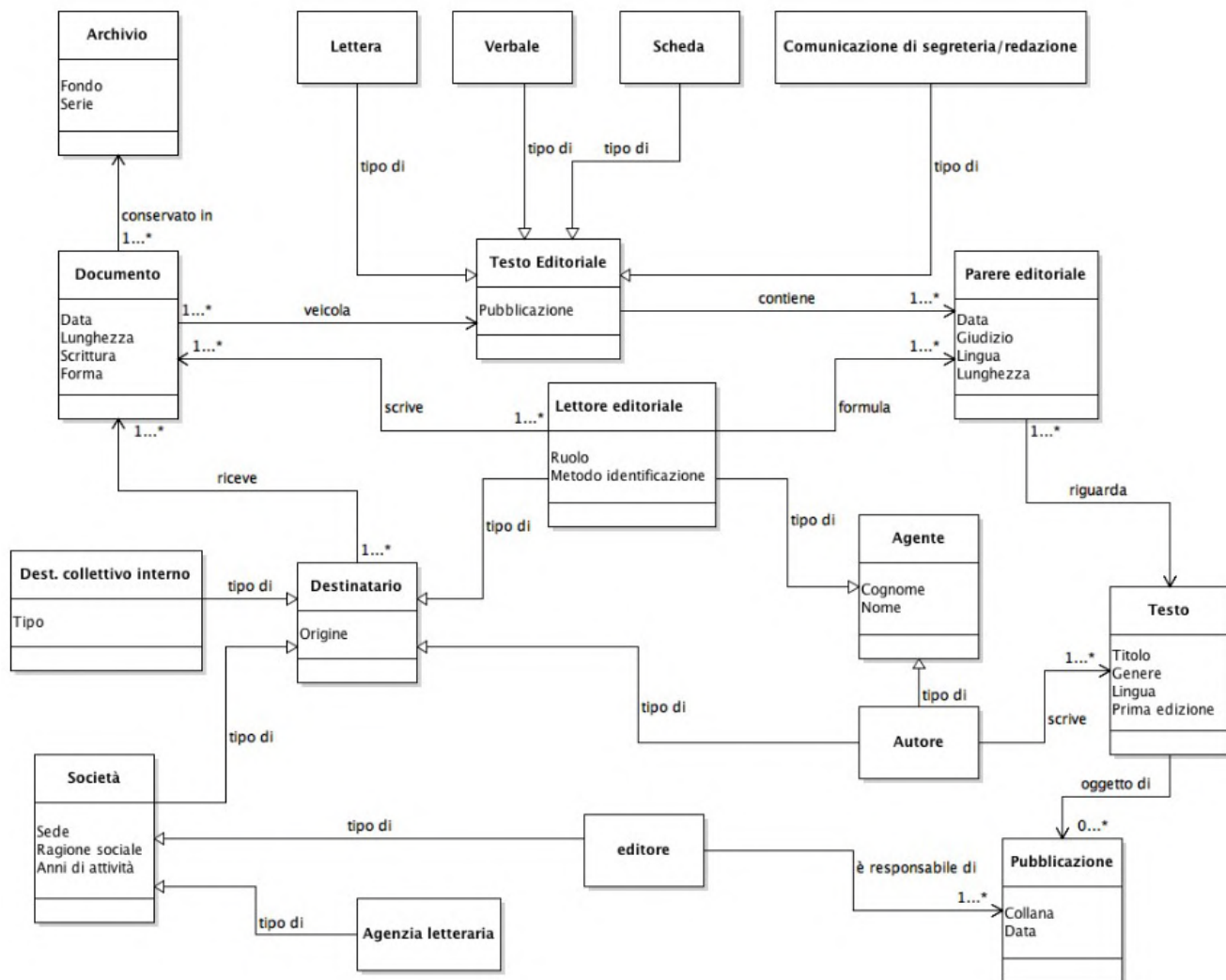


Figura 1

# eLearning the URLCoFi - Digital Didactics for Humanists

Armin Hoenen

Goethe University Frankfurt, DE, [hoenen@em.uni-frankfurt.de](mailto:hoenen@em.uni-frankfurt.de)

## ABSTRACT

In this article, an example of a flipped classroom course in searching content for Under-Resourced languages (or Low resource languages LRL) is being presented. While in the last century, availability of text in LRLs was limited, through the internet this gradually changes, but, given that the communities behind LRLs are often not economically prevalent, the content may be hard to locate among the millions of documents in bigger languages. How can one thus compile corpora in LRLs effectively from online sources and which factors play major roles in this are key questions to the current project.

## KEY WORDS

eLearning, Low Resource Languages, Under-Resourced Languages

## INTRODUCTION

Under Resources Language Content Finder (URLCoFi)<sup>1</sup> is the name of a project which is being supported and supervised at the Goethe University of Frankfurt by its central elearning authority [studiumdigitale](#). The goal of the project is to create a flipped classroom for digital linguists. The host, the [Institute of Empirical Linguistics](#) at Goethe University Frankfurt has a profound expertise in a wide range of smaller and historical languages such as Old High German, Irish, Lithuanian, Georgian and others used in the course.

*Net mining* is a metaphor for using the web as a ‘mine’. [1] introduce a special issue on compiling corpora from the web. Some time later [2] describes a methodology suitable for LRLs. [3] create BootCat, the to date most used tool for the purpose, it uses Bing.com and needs a list of input terms or URLs. Since not all content on the web is indexed by big search engines, other strategies may be necessary to compile large enough corpora for small enough languages, for examples see [4]–[6]. Specialized projects such as the An Crúbadán project [7] or meta collections of texts such as the Leipzig Corpora Collection [8] for LRLs exist but there are some legal barriers complicating availability. Despite all these efforts, it remains relatively hard to mine specialized corpora. Mining the net aiming at balanced corpora, cf. [BNC](#), has obvious and less obvious caveats (duplicats, main content detection [9] etc.) making it expensive and time-consuming. For many scientific works (e.g. research papers) a corpus is extremely helpful. For smaller languages of communities with a weaker financial backbone it is often impossible to finance a reference corpus, the Abkhaz national corpus [10] being one exception. This is regrettable because technologies such as machine translation rely on corpora, see [11].

## FRESH INTO THE DIGITAL - SURMOUNTABLES

Illustrating ever better data accessibility for languages, we compare two resources. When Greenberg wrote his famous book on universals in 1963 [13], he investigated 30 languages

---

<sup>1</sup> The ambiguity with URL (Uniform Resource Locator) is intentional since it helps memorization (as well as hopefully the pleasant evocation of coffee).

for convenience of availability of reliable information from printed grammars. The [WALS](#) [14] born in the digital age features roughly 2,700 but offers a most incomplete matrix. [Ethnologue](#) lists 7,100 recent (non-historical) languages for the world. Thus, smaller languages have come into reach, but content in some is quite hidden. The need to sift through big digital data masses is a key competence and surfaces very clearly for LRLs. We thus see a major need in teaching strategies/concepts that like distant reading [15] are developed in order to cope with big data (here is the buzz word we seeked to avoid but cannot hold back any longer). We suppose this question is more general although it might surface in LRL related issues since not sifting through masses of relevant data but finding relevant data among many irrelevant documents is important here.

A major problem is that there is no good way of evaluation for content from the internet [16]. How can one know how good the true recall is. Search engines have indexed only small parts of the internet and ways to manually reach the content beyond are limited by link-hopping capacities. This should not discourage the teaching process for this important competence. Using the existing knowledge of departments and bundling them into resources for recall estimation (verified URL-lists) and targetted manipulation of the web itself (placing content in a[n artificial] language) profiting from its fluidity are only two aspects this submission seeks to present as didactic concepts.

## **COURSE OUTLINE**

The course concept is a blended learning scenario and has an obligatory online phase which is realized as a LernBar [12] online lesson which precedes the presence phase where the learned is put into practice and discussed. In the first half of the course, participants will learn about existing text collections, search engines and query languages and crawling/scraping tools which might offer some initial text in their target LRL. In the second half, the weekly LernBar course will introduce a certain genre (e.g. newswire) or technical type of web document (e.g. presentations) and a certain LRL and combines them into an excersize to find documents fitting both the genre/filetype and the LRL. Results will then be analyzed in the presence phase with the ultimate goal to identify important quantitative (how many seeds, which frequency classes, which blacklisted terms etc.) and qualitative aspects (which terms reflect the socio-economic/political/geographical/medial setting and which linguistic implications does this have for good seed terms). Together participants and teacher formulate a tentative “best search practice guideline” for the current genre/LRL. At the end of the course, these guidelines are reviewed for more general patterns. We argue that such a course represents a possibilty/tool to dynamically improve the above mentioned poor availability of resources for LRLs.

## **CONCLUSION**

We have presented a concept of an eLearning course for LRLs which we think is generalizable in the context of teaching and expanding not only a key ability for humanities scholars in the age of big data, but also for coping with the recall evaluation problem. The proposal is in line with the purpose of universities as described by [17] in her conclusion as coordinated and acedamically orchestrated access to knowledge.

## **RINGRAZIAMENTI**

Vorrei ringraziare studiumdigitale.

## BIBLIOGRAPHY

- [1] A. Kilgarriff and G. Grefenstette, "Introduction to the special issue on the Web as corpus," *Comput. Linguist.*, vol. 29, no. 3, pp. 333–347, 2003.
- [2] S. Sharoff, "Creating general-purpose corpora using automated search engine queries," *WaCky Work. Pap. Web Corpus*, pp. 63–98, 2006.
- [3] M. Baroni and S. Bernardini, "BootCaT: Bootstrapping Corpora and Terms from the Web.," in *Proceedings of the LREC*, 2004.
- [4] A. Barbaresi, "Challenges in web corpus construction for low-resource languages in a post-BootCaT world," in *6th Language & Technology Conference, Less Resourced Languages special track*, 2013, pp. 69–73.
- [5] A. Barbaresi and K.-M. Würzner, "For a fistful of blogs: Discovery and comparative benchmarking of republishable German content," in *KONVENS 2014, NLP4CMC workshop*, 2014, pp. 2–10.
- [6] A. Barbaresi, "Finding viable seed URLs for web corpora: a scouting approach and comparative study of available sources," in *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*, 2014, pp. 1–8.
- [7] K. P. Scannell, "The Crúbadán Project: Corpus building for under-resourced languages," in *Building and Exploring Web Corpora: Proceedings of the 3rd Web as Corpus Workshop*, 2007, vol. 4, pp. 5–15.
- [8] C. Biemann, G. Heyer, U. Quasthoff, and M. Richter, "The Leipzig Corpora Collection-monolingual corpora of standard size," *Proc. Corpus Linguist.*, vol. 2007, 2007.
- [9] T. Gottron, "Evaluating content extraction on HTML documents," in *Proceedings of the 2nd International Conference on Internet Technologies and Applications (ITA'07)*, 2007, pp. 123–132.
- [10] P. Meurer, "The Abkhaz National Corpus.," in *LREC*, 2018.
- [11] S. Krauwer, "The basic language resource kit (BLARK) as the first milestone for the language resources roadmap," *Proc. SPECOM 2003*, pp. 8–15, 2003.
- [12] D. Weiß and M. Eichhorn, "LernBar—An Authoring Tool for Producing Web Content for Mobile Learning Using Responsive Design," presented at the Online Educa, Berlin, 2015.
- [13] J. H. Greenberg, *Universals of language*. Massachusetts Institute of Technology Press, 1963.
- [14] M. S. Dryer and M. Haspelmath, Eds., *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013.
- [15] F. Moretti, *Distant reading*. Verso Books, 2013.
- [16] L. Vaughan, "New measurements for search engine evaluation proposed and tested," *Inf. Process. Manag.*, vol. 40, no. 4, pp. 677–691, 2004.
- [17] R. Pace, *Digital Humanities, una prospettiva didattica*. Carocci Editore, 2015.

# 14-18 Documenti e immagini della Grande Guerra

Elisa Sciotti, Francesco Gandolfi

Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane e per le Informazioni Bibliografiche,  
Italia - elisa.sciotti@beniculturali.it; francesco.gandolfi@beniculturali.it

## PORTALE

*14-18 Documenti e immagini della Grande Guerra* è un archivio digitale coordinato dall'ICCU che nasce nel 2005 con l'obiettivo di riunire virtualmente il materiale bibliografico e archivistico del "Fondo Guerra"<sup>1</sup>, raccolto fin dall'inizio del conflitto dall'allora Comitato nazionale per la storia del Risorgimento italiano.

Attualmente la base dati, in costante incremento, rende disponibili oltre 620.000 contenuti digitali, forniti da 110 istituti culturali italiani tra archivi, biblioteche, musei, comuni e enti diversi. Una piccola parte di questi contenuti sono, inoltre, offerti da privati cittadini.

14-18.it contiene documenti pubblicati o prodotti in un periodo compreso tra il 1908 e il 1922, tale arco cronologico intende documentare il periodo del conflitto, considerando i prodromi e i successivi effetti della guerra.

Le risorse disponibili, suddivise in 21 differenti tipologie di materiale<sup>2</sup>, sono di natura pubblica e privata, fonti ufficiali e documenti detti "minori" o "effimeri".

L'archivio digitale contribuisce alla ricostruzione della memoria collettiva e offre un nuovo approccio alle fonti, consentendo di scoprire storie di interesse generale o personale.

Il portale si rivolge a un pubblico molto trasversale costituito sia da studiosi e ricercatori che indagano e analizzano fonti primarie, sia da persone appassionate di storia o interessate alla ricostruzione di eventi familiari. Il materiale reso disponibile può, inoltre, essere utilizzato per fini didattici e formativi<sup>3</sup>.

L'esplorazione dei materiali all'interno dell'archivio digitale avviene secondo criteri semplici: con una ricerca per parola chiave o attraverso la ricerca avanzata basata su di un set minimo di filtri. Il portale è dotato di un motore di indicizzazione e ricerca *Apache Solr* che garantisce elevate prestazioni e permette la navigazione a faccette consentendo all'utente di continuare a raffinare i risultati. I sistemi di *information retrieval* utilizzati e la ricerca full text in tutti i campi hanno dimostrato di soddisfare le richieste degli utenti.

## MODELLO DEI DATI

Il modello dati, comprende i principali campi descrittivi previsti dagli standard ISAD e ISBD in quanto tali normative sono le più idonee a descrivere la varietà dei documenti del portale, ma sono anche le più diffuse, in particolare tra archivi e biblioteche.

Gli istituti partner forniscono i metadati associati agli oggetti digitali nei formati originali da loro prodotti (xls, xlsx, csv, xml). I metadati, prima di essere pubblicati, vengono verificati e resi omogenei per ciascuna tipologia di materiale e spesso sono arricchiti con informazioni aggiuntive, sia a testo libero, sia sulla base di vocabolari controllati per migliorare il recupero delle risorse attraverso la ricerca.

---

<sup>1</sup> Per approfondimenti sulla nascita del portale si veda <http://www.14-18.it/progetto>

<sup>2</sup> Le 21 tipologie di materiale contenute nel portale sono organizzate nelle seguenti macrocategorie: Fotografie, Stampati, Documenti manoscritti, Grafica, Periodici, Cimeli, Monumenti e lapidi, RegISTRAZIONI sonore.

<sup>3</sup> Si veda, per esempio, la collaborazione a Generazioni connesse, il programma del MIUR a sostegno dell'educazione all'uso consapevole di Internet e dei media <http://www.generazioniconnesse.it/site/it/2018/02/02/la-grande-guerra/>. 14-18.it ha fornito uno schoolkit e un tutorial sul tema della Grande Guerra destinato ai docenti <http://www.14-18.it/bacheca/111>



“Monumenti e lapidi” è l’unica tipologia di materiale il cui modello dati è, invece, quello previsto dalla Scheda OA (Opere/oggetti d’Arte) elaborata dall’ICCD.

Il portale 14-18.it, attraverso opportune operazioni di mapping, ha dimostrato di soddisfare le esigenze di interoperabilità in uscita con altre piattaforme di aggregazione e contribuisce a Europeaana, l’infrastruttura europea per il patrimonio culturale digitale, esportando le proprie risorse nel formato Europeaana Data Model (EDM) via protocollo oai pmh.

## STRUTTURA

La struttura sviluppata dal portale per la rappresentazione descrittiva dei documenti digitali permette di associare i metadati a ciascuna immagine e offre, pertanto, la possibilità di descrivere il contenuto di singole parti che compongono un documento (es.: spoglio di periodici). Consente, inoltre, la descrizione sia di singoli documenti (unità documentaria e record bibliografico), sia di raccolte editoriali o composte da documenti di tipologia differente (unità archivistica) nelle quali alla scheda descrittiva d’insieme possono essere legate le schede descrittive specifiche di ciascun documento che è parte della raccolta.

Per il trattamento di “Periodici” e “Giornali di trincea” è prevista una struttura verticale ad albero che permette, una volta selezionato il titolo, di scegliere attraverso schermate successive l’annata e il fascicolo che si desidera consultare. La prima scheda visualizzata conterrà i dati relativi al periodico o al giornale di trincea, scorrendo le schede successive appariranno i dati relativi a ogni pagina/immagine visualizzata, compresi i titoli analitici, se inseriti.

Gli “Album fotografici” sono invece caratterizzati da una struttura orizzontale che mette in relazione le pagine dell’album con le singole fotografie contenute al suo interno. Ciascuna pagina dell’album e ogni singola fotografia sono dotate di una propria scheda descrittiva che può essere arricchita indipendentemente dalle altre e può, perciò, contenere informazioni specifiche. Tale approccio è stato particolarmente apprezzato da studiosi del settore, poiché consente di visualizzare l’album fotografico nella sua integrità originaria e di ricostruire i termini della sua formazione e della sua storia<sup>4</sup>.

## PROCEDURE D’IMPORTAZIONE

Le procedure d’importazione e pubblicazione degli oggetti digitali e dei metadati associati sono gestite *in-house* da *back end*, attraverso una piattaforma predisposta per il caricamento di file *xlsx* e *xml*. Il file Excel d’importazione comprende, oltre a tutti i campi descrittivi del modello dati, anche quelli tecnici per aggregare correttamente gli oggetti digitali in base alle specificità di ciascuna tipologia di materiale.

Nel caso di record catalogati nel Servizio Bibliotecario Nazionale (SBN), i metadati possono essere direttamente importati in formato JSON interrogando l’OPAC di Indice SBN con chiamata *http/https*, utilizzando il codice identificativo univoco (BID) di ogni record.

Il file *xml* relativo al caricamento dei metadati appartenenti alla tipologia “Monumenti e lapidi” è quello previsto per la Scheda OA in esportazione dal SIGEC WEB.

Gli oggetti digitali, organizzati in apposite cartelle all’interno del server di produzione, possono essere immagini in formato *jpeg* oppure risorse sonore in formato *mp3*.

Da *back end* è, inoltre, possibile cancellare e modificare i record importati grazie a una maschera di editing.

---

<sup>4</sup> Si veda “Si chiude un archivio, si apre un portale” di Erica Grossi <http://www.novecento.org/uso-pubblico-della-storia/si-chiude-un-archivio-si-apre-un-portale-gli-album-fotografici-della-guerra-nello-schermo-digitale-del-centenario-1893/>

14-18.it è dotato, infine, di un Content Management System (CMS) per la gestione di approfondimenti e comunicazioni, sitografia e raccolte di mostre virtuali.

## BIBLIOGRAFIA

- [1] De Capua, S., Martini, P., Metelli, P. *Europeana Collections 1914-1918 Ricordare la Prima Guerra Mondiale*. In: DigItalia: rivista del digitale nei beni culturali. ISSN: 1972-621X <http://digitalia.sbn.it/article/view/719/494>
- [2] Gioia, P. *Progetto Immagini della Grande Guerra*. In: DigItalia: rivista del digitale nei beni culturali. ISSN: 1972-621X <http://digitalia.sbn.it/article/view/342>
- [3] Gioia, P., Pizzo, M., Santiemma, A. *Ricordando la prima guerra mondiale*. In: DigItalia: rivista del digitale nei beni culturali. ISSN: 1972-621X <http://digitalia.sbn.it/article/view/553>
- [4] Martini, P. *La Grande Guerra nelle raccolte nazionali ed europee: materiali immagini e testimonianze*. In: DigItalia: rivista del digitale nei beni culturali. ISSN: 1972-621X <http://digitalia.sbn.it/article/view/1062/692>
- [5] Piccininno, M., Sciotti, E. *14-18, Documenti e immagini della Grande Guerra: un bilancio*. In: DigItalia: rivista del digitale nei beni culturali. ISSN: 1972-621X <http://digitalia.sbn.it/article/view/2040/1413>

# Il Laboratorio di Epigrafia Greca dell'Università Ca' Foscari. Una fucina didattica per l'epigrafia greca

Eloisa Paganoni, Stefania De Vido, Claudia Antonetti

Università Ca' Foscari Venezia – [eloisa.paganoni@unive.it](mailto:eloisa.paganoni@unive.it)

Università Ca' Foscari Venezia – [devido@unive.it](mailto:devido@unive.it)

Università Ca' Foscari Venezia – [cordinat@unive.it](mailto:cordinat@unive.it)

## ABSTRACT

Il contributo intende mettere in luce l'impatto didattico dei progetti a oggi in corso all'interno del Laboratorio di Epigrafia Greca del Dipartimento di Studi Umanistici di Ca' Foscari: *Axon. Silloge di iscrizioni greche* e *Venice Squeeze Project*. Grazie agli strumenti offerti dalle *digital humanities*, essi mirano ad avvicinare studenti di tutti i livelli e grande pubblico a una disciplina, l'epigrafia greca, e a un tipo di documento, l'iscrizione, ancora poco conosciuti.

## PAROLE CHIAVE

epigrafia greca; didattica; digital humanities; calchi epigrafici; public history

## PROPOSTA

La riforma universitaria del 1999 ha imposto di privilegiare le materie di base di ciascun settore scientifico disciplinare, determinando così la drastica riduzione degli insegnamenti specialistici, tra cui l'epigrafia greca. All'allontanamento da questa disciplina ha contribuito in misura non minore la scarsa conoscenza della lingua greca che costituisce prerequisito necessario all'approfondimento. Le conseguenze di questa situazione sono già in parte visibili: il mancato studio dell'epigrafia greca si traduce non solo in una diminuzione di quanti scelgono questa disciplina quale campo di specializzazione, ma anche in un uso limitato e spesso acritico del documento epigrafico all'interno di ricerche di taglio storico. La mancanza di formazione in ambito epigrafico e la scarsa conoscenza del greco antico costituiscono limiti che avranno conseguenze significative anche nelle competenze di esperti e funzionari preposti alla tutela e valorizzazione del patrimonio culturale.

Eppure, gli studenti di tutti i settori delle scienze dell'antichità (storia, filologia e letteratura, archeologia) manifestano un crescente interesse e accolgono favorevolmente l'offerta di corsi anche di livello avanzato e di iniziative didattiche, come visite a collezioni epigrafiche o viaggi di studio, che li mettano a diretto contatto con questa tipologia documentaria. Alla luce di questo, Claudia Antonetti e Stefania de Vido, che guidano il Laboratorio di Epigrafia Greca dell'Università Ca' Foscari<sup>5</sup>, promuovono l'insegnamento dell'epigrafia greca tramite diverse iniziative tra le quali spiccano due progetti dal forte impatto didattico che guardano alla realtà delle *digital humanities*.

Il primo è *Axon. Silloge di iscrizioni greche*<sup>6</sup>: nato nel 2014 grazie a un finanziamento d'Ateneo (PRA 2013), Axon ha portato alla creazione del primo database italiano di iscrizioni storiche greche, che raccoglie i principali documenti epigrafici dall'epoca arcaica a quella ellenistica, con schede preparate da oltre cento studiosi attivi in Italia e all'estero, liberamente consultabili grazie a una maschera d'interrogazione semplice e versatile. La seconda iniziativa, coordinata da Claudia Antonetti, è il *Venice Squeeze Project*<sup>7</sup>: nato anch'esso

---

<sup>5</sup> <http://virgo.unive.it/archeolab/index.php?it/264/laboratorio-di-epigrafia-greca>.

<sup>6</sup> <http://virgo.unive.it/venicepigraphy/axon/public/>.

<sup>7</sup> <http://mizar.unive.it/venicesqueeze/public/frontend/collezione>.

grazie a un finanziamento d'Ateneo (SAL 2016), ha preso avvio nel settembre del 2017 e mira alla digitalizzazione della collezione di calchi di iscrizioni greche del Laboratorio di Epigrafia Greca, composta da 605 documenti. Il progetto si sviluppa in collaborazione con il Laboratorio HiSoMA dell'Université Lumière Lyon 2 e con l'École française d'Athènes attualmente impegnati nella realizzazione di *E-Stampages*<sup>8</sup>, una piattaforma digitale open-access che ospiterà la collezione veneziana e una cospicua parte delle collezioni di calchi epigrafici dei due istituti francesi, rimanendo aperta a future implementazioni da parte di quanti vorranno pubblicare in essa le proprie collezioni.

Delle questioni specifiche che la realizzazione di questi progetti ha sollevato (prime tra tutte la selezione e organizzazione dei metadati e la costituzione di vocabolari controllati) e delle soluzioni via via adottate si è già discusso in altre sedi<sup>9</sup>; vogliamo qui invece considerare l'aspetto didattico di entrambe le iniziative. *Axon* è stato concepito per rendere accessibile il documento epigrafico anche agli studenti che si accostano per la prima volta all'epigrafia greca e a un pubblico di non specialisti, superando il primo ostacolo apparente alla fruizione e alla comprensione, la lingua. Ciascuna iscrizione, presentata in lingua originale secondo rigorosi criteri filologici, è infatti accompagnata dalla traduzione del testo, che rappresenta la 'porta di accesso' al documento. Lo studente può così procedere gradatamente alla 'scoperta' del documento, consultando la scheda che raccoglie tutte le informazioni ad esso relative, dal supporto al luogo di ritrovamento e conservazione, dalla datazione alla tipologia del testo. Chi conosce il greco può trovare di grande utilità sia la sezione dedicata alle notazioni paleografiche e linguistiche, sia il lemma genetico, che elenca le edizioni ponendone in luce le dipendenze e che costituisce per certi versi il fulcro metodologico della scheda, sia, infine, il testo e l'apparato critico. Attraverso la consultazione delle schede lo studente prende consapevolezza dei dati necessari a descrivere un'iscrizione e della loro organizzazione logica e conseguente; contemporaneamente egli sviluppa familiarità con il lessico specifico della disciplina. Il passo successivo è la contestualizzazione del documento grazie al commento, il quale evidenzia l'apporto dell'iscrizione alla comprensione del contesto storico. La consultazione delle schede *Axon* non richiede specifiche conoscenze, ma contribuisce a svilupparle, con la conseguenza di avvicinare gli studenti alla disciplina, studenti che talvolta mettono in pratica quanto acquisito nel proprio percorso formativo divenendo a loro volta autori di nuove schede *Axon*.

Chi si occupi di iscrizioni sa che l'autopsia del documento è fondamentale, ma non sempre possibile. Per ovviare a questa situazione fisiologica, gli studiosi cercano di ottenere riproduzioni il più possibile fedeli al testo attraverso diverse modalità, tra cui la più efficace è il calco cartaceo, che offre una copia del testo identica all'originale sulla quale verificare letture e possibili integrazioni. Il calco è anche un utile strumento didattico: tramite esso infatti gli studenti possono entrare in contatto con il documento originale, imparando a distinguere le lettere anche nelle loro forme meno usuali e prendendo consapevolezza delle peculiarità paleografiche sviluppatasi nelle diverse regioni e nelle diverse epoche. È però vero che poche sono le istituzioni che dispongono di collezioni di calchi epigrafici e perciò non stupisce il fiorire di collezioni digitali di questo genere di documenti<sup>10</sup>. Ponendosi in questo solco, *Venice Squeeze Project* ed *E-Stampages* si caratterizzano per il tipo di riproduzione del calco che essi intendono offrire: se in genere le collezioni digitali presentano una riproduzione statica dell'oggetto, sia essa bi- o tridimensionale, il progetto veneziano, insieme con i partner francesi, vuole proporre una ricostruzione 3D che l'utente possa muovere, inclinare e ruotare, replicando di fatto i gesti che gli studiosi compiono per leggere i calchi. Ottenuta con il

---

<sup>8</sup> <https://www.efa.gr/index.php/en/ressources-documentaires/les-archives/archives-estampages/le-programme-e-stampages>.

<sup>9</sup> De Vido, Matijasic, Palazzo 2017; Antonetti, De Vido 2017.

<sup>10</sup> Per una selezione, Antonetti, De Vido 2017, 492.

software open source DEA, messo a punto da Angelos Barmpoutis e Eleni Bozia (University of Florida) nell'ambito del *Digital Epigraphy and Archaeology Project*<sup>11</sup>, questa riproduzione ha un forte impatto didattico, in quanto consegna nelle mani degli studenti uno dei principali strumenti di lavoro degli epigrafisti.

Nell'insieme, i progetti promossi dal Laboratorio di Epigrafia Greca possono essere considerati due iniziative concrete che, con modalità differenti, avvicinano gli studenti all'*ars epigraphica* e alla consultazione del documento originale avvalendosi dei metodi sviluppati dalle digital humanities.

## BIBLIOGRAFIA

- [1] Antonetti, C., De Vido, S. 2017. *Digital Epigraphy at the Greek Epigraphy Laboratory, Ca' Foscari University of Venice*. In: *Historika* 7, 491-502. DOI= 10.13135/2039-4985/2608.
- [2] Barmpoutis, A. Bozia, E., Wagman, R.S. 2010. *A Novel Framework for 3D Reconstruction and Analysis of Ancient Inscriptions*. In: *Machine Vision and Applications*. DOI= 10.1007/s00138-009-0198-7.
- [3] De Vido, S., Matijasic, I., Palazzo, S. 2017. *AXON. A Database for Greek Historical Inscriptions. Digital and Traditional Epigraphy in Context*. In: *Digital and Traditional Epigraphy in Context. Proceedings of the EAGLE 2016 International Conference*. Roma, 57-65. ISBN 978-88-9377-021-7.  
[http://www.editricesapienza.it/sites/default/files/5533\\_EAGLE\\_2016\\_International\\_Conference.pdf](http://www.editricesapienza.it/sites/default/files/5533_EAGLE_2016_International_Conference.pdf). DOI= 10.13133/9788893770271.
- [4] Levivier, A., Leblanc, E., Brunet, M. 2016. *E-Stampages. Archivage et publication en ligne d'une ectyothèque d'inscriptions grecques*. In: *Nouvelles de l'archéologie* 145.  
<http://journals.openedition.org/nda/3801>. DOI= 10.4000/nda.3801.
- [5] Tracy, S.V. 2010. *What are Squeezes and How are They Used?*. In: *The Institute Letter* Fall 2010. <https://www.ias.edu/ideas/2010/tracy-ias-squeeze>.

---

<sup>11</sup> <http://www.digitalepigraphy.org/>.

# Complessità della codifica ed ergonomia strumentale nel contesto XML-TEI: dove siamo? (Bilancio a partire da un nuovo progetto di edizione digitale medievale)

Marta Materni

IF Marie Curie, LUHCIE, UGA-Université Grenoble Alpes, [marta.materni@univ-grenoble-alpes.fr](mailto:marta.materni@univ-grenoble-alpes.fr)  
[marta.materni@gmail.com](mailto:marta.materni@gmail.com)

## ABSTRACT

Il contributo, nato dall'esperienza maturata nell'ambito di un nuovo progetto di edizione digitale medievale, intende fare il punto sulla questione, ancora problematica, della produzione di file XML-TEI. Gli strumenti, ancora oggi come un decennio fa quando Peter Robinson aveva denunciato il problema, sono ben lontani dall'essere ergonomici ed efficaci. La questione è spinosa in quanto la lacuna strumentale incide sia sulla diffusione della pratica editoriale digitale – la mancanza di strumenti non è il solo fattore che ha frenato quello che, stando all'idea di una possibile *rivoluzione*, doveva conseguentemente presentarsi come una crescita esponenziale delle edizioni digitali realizzate secondo principi scientifici; sia, ancor peggio, e assai più direttamente, sulla qualità della codifica proposta. Il rischio che si sta correndo è quello di continuare a produrre “codifiche semplici con strumenti complessi” quando ormai la quotidianità filo-digitale dovrebbe essere la produzione di “codifiche complesse con strumenti semplici”. Il progetto di edizione messo in opera ha implicato la creazione di un prototipo di editor XML-TEI capace di generare il codice in modo automatico a partire da una sintassi simbolica (sul modello di Markdown) applicata a un semplice file .TXT associato con un file .CSV, con la possibilità di ottenere inoltre molteplici codifiche a partire dallo stesso file modificando le corrispondenze fra simboli e tag XML nel file .CSV. È stato possibile in questo modo realizzare una codifica complessa secondo un modello di testo concepito come un database di parole, pronto a essere utilizzato immediatamente per altre analisi al di là degli obiettivi specifici di edizione del contesto progettuale in cui il file è stato prodotto. Si precisa in questo modo, anche, il ruolo del “filologo digitale”, responsabile della diffusione all'interno (e all'esterno) della comunità accademica di un testo filologicamente **E** informaticamente corretto, “pronto all'uso”, secondo un modello di capitalizzazione progressiva del lavoro di analisi.

## PAROLE CHIAVE

Edizione digitale, TEI, editor TEI, modellizzazione della codifica

## TESTO

Nel 2005, in un articolo ancora (purtroppo) di attualità e che offre numerosi spunti di riflessione, Peter Robinson lanciava un grido di allarme: «We need some things we do not yet have: software that does not exist and established online publication systems that have yet to be created. Let us not wait too long for these» (Robinson 2005, § 31). La letteratura grigia lascia spesso trasparire del disagio al riguardo, e nemmeno tanto velato. Tara Andrews, autrice di una soluzione parzialmente analoga a quella qui proposta, TEI Markup, è piuttosto efficace nella descrizione:

TEI XML is a wonderful thing. [...] The problem is the transcription itself. When I am transcribing a manuscript [...] I do not want to be switching back and forth between keyboard layouts in order to type `<tag attr="attr">arrow-arrow-arrow-arrow-arrow</tag>` every six seconds. It's prone to typo, it's astonishingly slow, and it makes my wrists hurt just to think about it. (<https://metacpan.org/pod/Text::TEI::Markup>)

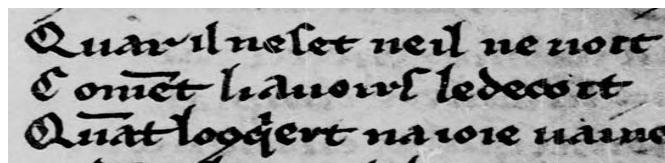
Recentemente Roberto Rosselli del Turco ha ripreso la questione in un articolo di sintesi del 2016 che prende in considerazione i due poli della problematica: produzione e visualizzazione di un'edizione digitale. L'argomento è talmente vasto che in questo contesto vorrei soffermarmi solo sul polo produzione, che si rivela strettamente connesso alla questione della complessità del modello di codifica proposto: “complessità di codifica di una codifica complessa” potrebbe essere la definizione sintetica del problema. La conclusione di Rosselli del Turco un decennio dopo la denuncia di Robinson costituisce la base di partenza per la riflessione:

Robinson's remark about the lack of easy-to-use production tools is unfortunately still valid: there is no software tool nor suite of tools that allows a scholar to produce a full digital edition, be it image-based with a diplomatic text or a critical edition, in a way comparable to how printed editions are prepared (Rosselli del Turco 2016, § 16)

Il problema strumentale è ovviamente scientificamente piuttosto grave in quanto rischia di condizionare in modo *inaccettabile* le nostre proposizioni teoriche di modellizzazione della codifica testuale, il che significa in ultima analisi una prospettiva in cui lo strumento tecnico ha il potere di influenzare le nostre decisioni intellettuali riguardanti la rappresentazione del pensiero. Ancora una frase di Robinson a farci da guida:

A well-made electronic scholarly edition will be built on encoding of great complexity and richness. As well as free text searching, efficient search systems can make use of this encoding to enable sophisticated search, going considerably beyond the standard word and phrase search. (Robinson 2004)

Analisi testuali complesse sono possibili solo a partire da una codifica testuale complessa: altrimenti, a che pro ricorrere alla dimensione digitale quando quella stampata è già così roduta ed efficiente rispetto a una certa ben radicata visione dell'edizione? Gli strumenti proposti ufficialmente, come *Oxygen*, con la loro farraginosità di produzione, ci inchiodano di fatto a codifiche al limite talvolta della banalità. Di fronte alla ricchezza di informazioni estrapolabili da un verso in antico francese tramandato da un manoscritto medievale – sistema abbreviativo, stadio evolutivo della lingua, grafie regionali, segmentazione della scrittura, forme paleografiche ecc.



quanti dati perdiamo proponendo una codifica di questo tipo, purtroppo piuttosto diffusa?

```
<l>Quar il ne set ne il ne voit</l>  
<l>Come<ex>n</ex>t li avours le deçoit</l>  
<l>Qua<ex>n</ex>t lo <ex>con</ex>q<ex>ui</ex>ert n'a joie vaine</l>
```

Quanto possiamo invece guadagnare in potenziale di analisi con una codifica di quest'altro tipo, che trasforma il testo in un database di parole, ciascuna identificabile e quindi analizzabile e quindi manipolabile? Limitiamoci al terzo verso:

```

<l n="3" xml:id="msl3">
  <w xml:id="msl3w1">Qua<expan corresp="#abb-tild-nas"><ex>n</ex></expan>t</w>
  <w xml:id="msl3w2" ana="#aggl-s">lo</w>
  <w xml:id="msl3w3"><expan corresp="#abb-tir-9"><ex>con</ex></expan>
    <expan corresp="#l-sup-q">q<ex>u</ex>i</expan>ert</w>
  <w xml:id="msl3w4" ana="#elis">n</w>
  <w xml:id="msl3w5" ana="aggl-s-unc">a</w>
  <w xml:id="msl3w6"><c ana="#lram-cons">i</c>oie</w>
  <w xml:id="msl3w7"><c ana="#lram-cons">u</c>aïne</w>
</l>

```

Possiamo in questo modo ottenere una visualizzazione diplomatica in modalità “trascrizione”:

Quant loconquiert na-ioie uaine

e in modalità “interpretativa”:

Quant lo conquiert n'a joie vaine

Si tratta del modello di codifica proposto per un recente progetto di edizione diplomatica digitale dei testimoni di un romanzo francese del XII sec., e pensato come una sorta di modello di codifica proiettato sul futuro, al di là degli obiettivi contingenti legati al progetto specifico: il testo, cristallizzato nella forma documentaria che si offre ai nostri occhi, viene scomposto nei suoi elementi minimali portatori di significato – la parola (un paleografo, interessato all’immagine del testo, probabilmente considererà come unità minima il carattere) – e ricostruito sotto forma di un file testuale codificato che vuole rappresentare un avatar digitale, cioè *formalizzato*, della realtà manoscritta. Così preparato, il file XML-TEI ha soddisfatto gli obiettivi immediati del progetto – la realizzazione di un’edizione diplomatica visualizzabile in modalità trascrizione grafematica e in modalità interpretativa modernizzata – , offrendosi al contempo come prodotto “già pronto” per essere sottoposto a ulteriori analisi: lemmatizzazione e analisi morfo-sintattica, utilizzo del *Double end-point attached method* per la creazione di un’edizione critica, analisi sintattiche, retoriche ecc. facilmente codificabili a questo punto attraverso uno stand-off markup dal momento che ciascuna parola è associata con un ID. Si tratta di applicare il principio metodologico dei “nani sulle spalle dei giganti”, senza naturalmente, fuor di metafora, attribuire alcun valore di “nanismo” o “gigantismo” scientifico ai due momenti; o, in altri termini, di capitalizzare al massimo il lavoro, cioè il tempo della ricerca. La posizione pragmatica, direi quasi imprenditoriale, rispetto alla codifica, sostenuta con forza per es. da Elena Pierazzo:

We all know how important economic considerations are in our decision-making processes; almost all of our research projects are funded for a specific time-span and budget, and so it is fundamental to ensure that the transcription (and encoding) is feasible within this lifetime. The decision whether or not to encode a specific feature will be heavily determined by the cost of encoding it, and it would be naive not to admit this. Economic considerations may then be used as the pragmatic limits of the level of transcription we are looking for, in the same way that the limits of the typography worked for print-based publications. [...] How far should we go in considering the needs of the Others? [...] if it is matter of considering the needs of possible future scholars in other disciplines and providing special markup for them, the temptation is to say: ‘not far’. There is, in fact, a serious risk of wasting precious project time here: it is very difficult to guess other scholars’ needs or principles which are, potentially, infinite – as infinite as the set of ‘facts’ that can be derived from the document. (Pierazzo 2011, p. 469, p. 471)

è esattamente il contrario di quanto si vuole qui proporre, e cercare di perseguire. Se tempo e denaro influenzano ovviamente le nostre attività, l’obiettivo scientifico non potrà essere tuttavia quello di assecondare le “leggi del mercato” quanto piuttosto di alleggerire il peso del loro gioco attraverso un percorso di sperimentazione e ricerca che permetta di “abbassare i



costi e ottimizzare i tempi”, per continuare a utilizzare questa metafora mercantile, ottenendo ciò che viene reputato scientificamente corretto e necessario e non “economico”.

Inoltre, la preoccupazione riguardo «how far should we go in considering the needs of the Others» potrebbe essere in realtà una falsa preoccupazione derivante dalla tendenza a cumulare nello stesso file gli strati di codifica descrittiva e gli strati di codifica interpretativa. Piuttosto che prevedere uno «special markup» per ciascuna, possibile disciplina, in quanto filologi creatori di un’edizione digitale dovremmo in realtà proporre una sorta di “markup passepartout” che descriva ma soprattutto formalizzi il testo, al quale poi le varie discipline applicheranno la loro griglia, e codifica, interpretativa. Lo stand-off markup è dal mio punto di vista la risposta per eccellenza, risposta che presuppone, per essere praticata correntemente, la formalizzazione del testo come agglomerato di elementi minimi identificati in modo univoco e utilizzabili quindi come *anchor*.

La questione diventa dunque: come creare in tempi ragionevoli e in una modalità scevra da errori ricorrenti un prodotto complesso e utile, a me e agli altri? Non certo attraverso editor XML “prestati” al mondo TEI, bensì attraverso editor XML-TEI *specificatamente* concepiti per la codifica testuale, e, in più, attraverso l’automatizzazione massimale della produzione dei tag XML. Più il processo è automatizzato, più l’errore della digitazione è ridotto, più la complessità della codifica è incrementata.

Se, ed è evidente, il processo di produzione con strumenti come Oxygen è farraginoso e lento, parimenti non reputo che la soluzione del problema possa passare attraverso piattaforme online come T-Pen o la nuova nata TextualCommunity, che semplicemente spostano il problema dal l’ambiente di lavoro “in locale” a quello “in linea”, incatenandoci fra l’altro a un utilizzo online che rappresenta una vera e propria schiavitù oltre a soffrire di una certa fragilità tecnologica. Né attraverso strumenti che nascondono il codice e che rappresentano un vero e proprio ossimoro teorico rispetto all’idea della codifica testuale: se il primo insegnamento rispetto al senso profondo della TEI insiste sulla differenza fra un sistema WYSIWYG e uno WYSIWYM, utilizzando uno strumento che cela ai nostri occhi il codice XML-TEI noi stiamo esattamente producendo un codice WYSIWYM con uno strumento WYSIWYG, stiamo cioè andando contro la stessa filosofia di cui dovremmo essere divulgatori.

La leggibilità umana del file XML non può a mio avviso rappresentare un principio guida della codifica, a meno di rinunciare, in nome di questa leggibilità, alla possibilità di analisi radicalmente alternative al paradigma editoriale cartaceo. Quello che propongo in alternativa è un processo di produzione che passa attraverso dei file testuali “di mediazione”, basati su un linguaggio simbolico ispirato alla sintassi Markdown HTML et alle *entities* XML (&...;). Si tratta della metodologia sperimentata nella realizzazione di questo progetto di edizione digitale e concretizzata nella creazione di un editor-prototipo: TEI-med.it, realizzato in Python. I simboli, applicati a un semplice file .TXT realizzabile con qualsiasi editor (Word *ovviamente* escluso), sono associati, in un file .CSV esterno, liberamente modificabile dall’utente, ai tag XML, con il duplice vantaggio, oltre alla rapidità di realizzazione, di poter proporre più soluzioni di codifica a partire dallo stesso .TXT, semplicemente associando un differente .CSV.

Riprendendo il verso di cui si è proposto il codice XML precedentemente, quel codice è stato ottenuto a partire da questa stringa testuale:

```
Qua&n;t +lo &9;&q;ert 'n +_a *ioie *uaine
```

Il linguaggio TEI-Med – che è solamente, sottolineo, un linguaggio di mediazione, strumentale, non un markup alternativo a TEI – si avvale di tre categorie di simboli:

1. simboli ASCII associati a una lettera o a una parola per produrre intere stringhe di codice o per generare un attributo:

es. \* associato a i, \*i, produce:

```
<c ana="#lram-cons">u</c>
```

o, volendo una codifica più classica:

```
<choice><orig>u</orig><reg>v</reg></choice>
```

2. *entities* fisse, corrispondenti a un singolo tag o a una stringa intera di codice, assai utile per le abbreviazioni, la struttura è &...;

es. &n; produce:

```
<expan corresp="#abb-tild-nas"><ex>n</ex></expan>
```

o, volendo:

```
<choice><abbr><am><g ref="tild-nas"/></am>
<expan><ex>n</ex></expan></choice>
```

3. *entities* con argomenti variabili, la struttura è &...;(\$)

es. &sub-ex-int;(parla,parola), corrisponde a:

```
<sub><del                                rend="erasure"><w
xml:id="ms11w1">parla</w></del>
<add                                place="interline"><w
xml:id="ms11w2">parola</w></del></sub>
```

Le categorie 2 e 3 sono liberamente modificabili e incrementabili dall'utente finale attraverso il .CSV. Linee <l>, parole <w>, @xml:id e @n vengono prodotti automaticamente dall'editor senza alcun intervento da parte dell'operatore sul .TXT.

Il sistema ha permesso di realizzare la trascrizione completa e la codifica di 6 manoscritti di un testo di 13.500 versi (all'inizio del progetto ne erano stati previsti 4) nello spazio di circa 18 mesi. Le edizioni sono in corso di pubblicazione sul sito *DigiFlorimont* (<http://digiflorimont.huma-num.fr/>).

## BIBLIOGRAFIA

- [1] Pierazzo, E. 2011. *A Rationale of Digital Documentary Editions*. In: *Literary and Linguistic Computing*, vol. 26.4: 463-477.
- [2] Robinson, P. 2004. *Where We Are with Electronic Scholarly Editions, and Where We Want to Be*. <http://computerphilologie.uni-muenchen.de/jg03/robinson.html>
- [3] Robinson, P. 2005, *Current Issues in Making Digital Editions of Medieval Texts—Or, Do Electronic Scholarly Editions Have a Future?*, In: *Digital Medievalist*, vol. 1.1. <http://www.digitalmedievalist.org/journal/1.1/robinson>  
DOI=<http://doi.org/10.16995/dm.8>
- [4] Rosselli del Turco, R. 2016. *The Battle We Forgot to Fight: Should We Make a Case for Digital Editions?*. In: *Digital Scholarly Editing: Theories and Practices*, Cambridge, Open Book Publisher, <http://books.openedition.org/obp/3423>>. ISBN: 9782821884007.
- [5] Sahle, P. 2016. *What is a Scholarly Digital Edition?*. In: *Digital Scholarly Editing: Theories and Practices*, Cambridge, Open Book Publisher, <http://books.openedition.org/obp/3423>>. ISBN: 9782821884007.

- [6] Stolz, M. 2017. Copying, Emergence and Digital Reproduction. Transforming Medieval Manuscript Culture into an Electronic Edition. In: *Digital Philology*, vol. 6.2: 257-87.

# “Osservate, leggete con me”<sup>1</sup>, Risorse LOD per la storia del melodramma: una prospettiva funzionale di rappresentazione

Paolo Bonora<sup>1</sup>, Angelo Pompilio<sup>2</sup>

<sup>1</sup> Dipartimento di Beni Culturali – Università di Bologna, Italia, paolo.bonora@unibo.it

<sup>2</sup> Dipartimento di Beni Culturali – Università di Bologna, Italia, angelo.pompilio@unibo.it

## ABSTRACT

L'intervento presenta i primi risultati del progetto *Corago LOD*, promosso dal Dipartimento di Beni Culturali dell'Università di Bologna, che mira alla sperimentazione delle tecnologie del Semantic Web applicate ad archivi digitali in ambito musicale. Il progetto si è posto tre obiettivi principali nell'ambito della sperimentazione delle tecnologie semantiche in ambito musicale: verificare gli esiti della trasposizione in formato RDF dei contenuti utilizzando ontologie di riferimento per il CH [6]; dimostrare che la base di conoscenza così prodotta fosse in grado di sostenere un'applicazione di navigazione efficace dal punto di vista dei fruitori; che il formato RDF rendesse possibile l'integrazione di contenuti ottenuti da fonti informative esterne secondo i principi Linked Data [1].

Il sistema informativo *Corago* è stato progettato come strumento per la ricerca storica e documentale sulla produzione del melodramma italiano, dalle sue origini (1600) ai primi decenni del secolo scorso. Nella tradizione delle arti performative l'opera lirica è senz'altro lo spettacolo che presenta la maggiore complessità, data da 'testi' di natura diversa (letteraria, musicale, coreografica, visiva), ma necessariamente coordinati. Di conseguenza la documentazione dello spettacolo operistico è caratterizzata da un'intrinseca multiformità: documenti verbali, musicali, iconografici, sonori, visivi. Potenzialmente *Corago* può contenere tutte le informazioni storiche e tutti i documenti sul teatro d'opera riprodotti in formato digitale [27]. Le dimensioni vaste della produzione operistica e la quantità sterminata di documenti hanno imposto però di delimitare l'ambito di intervento: al momento *Corago* si propone come strumento di supporto alla ricerca sulla storia materiale dello spettacolo operistico e allo studio della tradizione del testo letterario del melodramma italiano. In questa prospettiva di studi, si è concentrata l'attività di ricerca documentale su un solo tipo di documento: il libretto d'opera, fonte primaria insostituibile, che contiene il testo letterario e una ricca messe di informazioni sugli spettacoli rappresentati [8]. Il sistema informativo è articolato in tre contenitori: due di entità immateriali, il **Repertorio** delle opere prodotte e la **Cronologia degli spettacoli**, uno di oggetti materiali, l'**Archivio digitale dei libretti**, riprodotti integralmente e consultabili liberamente online.

Il modello concettuale *Corago* ed il processo di catalogazione che ha portato alla costruzione dell'attuale sistema utilizzato per la catalogazione e la consultazione è una implementazione del modello FRBR [19], con elementi derivati da FRAD<sup>2</sup>, che, pur mantenendosi coerente nell'impianto complessivo, ne rappresenta una significativa specializzazione.

Per l'esplicitazione della semantica del modello originale e la trasposizione dei contenuti in formato RDF sono state adottate le due ontologie di riferimento degli ambiti museali e bibliografico: CIDOC Conceptual Reference Model (CIDOC CRM) [15] e FRBR Object Oriented (FRBRoo) [12].

---

<sup>1</sup> Dall'aria di Leporello «Madamina, il catalogo è questo» (atto I, scena 5) del *Don Giovanni* di W.A. Mozart.

<sup>2</sup> FRAD: Functional Requirements for Authority Data è il modello concettuale di riferimento proposto dall'IFLA per la descrizione delle entità censite nell'authority file. Nasce come complemento a FRBR per la definizione di soggetti, thesauri, tassonomie e relazioni a supporto delle entità in esso definite [20].

Nell'insieme il processo di esplicitazione della semantica di modello ha permesso di rappresentare l'intero patrimonio informativo dell'archivio originale richiedendo solo minime estensioni e specializzazioni delle due ontologie di riferimento, in larga parte dovute alle specificità del dominio. Contestualmente, le dimensioni di spazio e tempo (ambientazioni, date e luoghi di rappresentazione o edizione, etc.) sono state ricondotte a coordinate comuni per rendere possibile il loro utilizzo come criteri universali di accesso ai contenuti<sup>3</sup>.

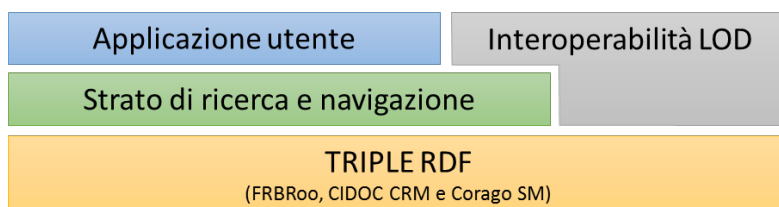


Figura 1

Per rendere efficace l'accesso ai contenuti informativi è stato necessario identificare una modalità di rappresentazione dell'informazione più vicina alla prospettiva dell'utente finale (figura 1). A tale scopo è stata realizzata una specifica semantica di navigazione che concettualmente si colloca in posizione mediana tra l'utente e la semantica di modello, qualcosa di simile alle lenti semantiche [9] [27], ma con la peculiarità di considerare questo strato intermedio di navigazione come parte integrante della base di conoscenza.

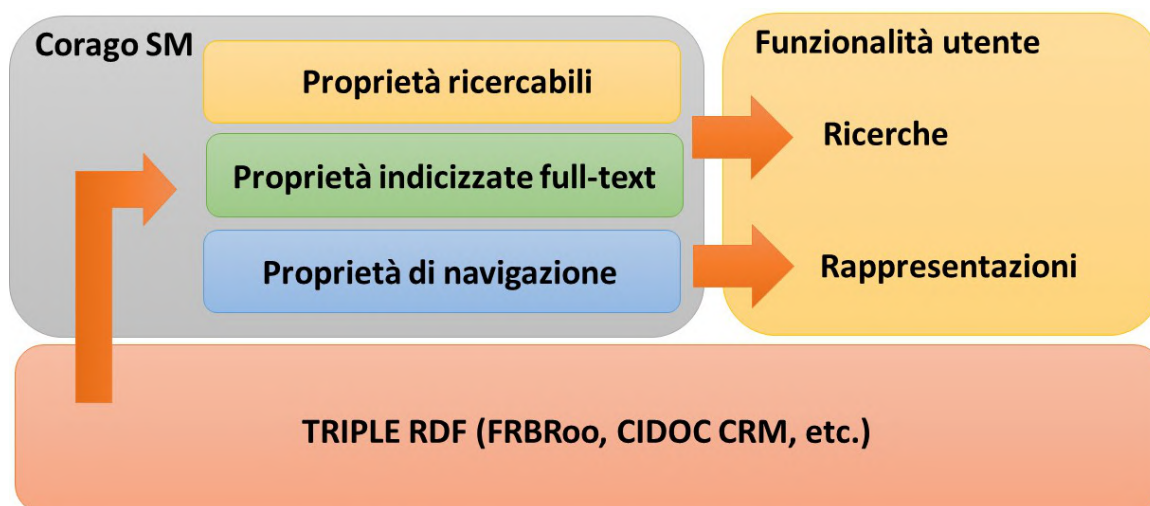


Figura 2

L'applicazione per la ricerca e la navigazione dei contenuti si avvale quindi di uno strato di astrazione definito dall'ontologia di dominio *Corago Semantic Model* in cui confluiscono, oltre alle necessarie estensioni delle due ontologie di riferimento, anche la definizione delle proprietà utilizzabili per ricercare e rappresentare agli utenti le tre entità principali del dominio: opere, eventi e documenti<sup>4</sup> (figura 2). Per la costruzione dell'indice full text sono

<sup>3</sup> L'informazione relativa alla dimensione spaziale è stata arricchita sfruttando l'allineamento con due basi di conoscenza: DBpedia e LinkedGeoData. La prima è stata utilizzata per l'analisi dei toponimi e la relativa disambiguazione, la seconda per l'estrazione delle informazioni geografiche.

<sup>4</sup> Le proprietà che definiscono i criteri di ricerca sono disegnate in modo da produrre come risultati entità di tipo *F1\_Work* (opere), *F31\_Performance* (Eventi), *F3\_Manifestation\_Product\_Type* (Documenti) che sono considerate le entità principali del dominio e i punti di accesso privilegiati da cui l'utente può procedere con la

state definite specifiche proprietà che identificano gli elementi informativi più significativi di ciascuna tipologia di risorsa per questo tipo di ricerca. Il risultato ottenuto è la possibilità per gli utenti di accedere ai contenuti attraverso chiavi testuali ed ottenere per ciascun risultato le relazioni con gli attributi contenenti il testo ricercato. In questo modo viene offerta all'utente piena evidenza del risultato ottenuto nonché la possibilità di approfondire la ricerca navigando il grafo attraverso le relazioni proposte dal sistema. Il disaccoppiamento tra dato RDF e la sua rappresentazione offre la possibilità di rendere omogenea in *Corago LOD* la fruizione di risorse provenienti da altre basi di conoscenza sebbene definite con ontologie diverse da quelle di riferimento. Nel complesso la soluzione elaborata affianca alla analiticità nella definizione della semantica dei contenuti offerta dalle due ontologie di riferimento, prerequisito per una piena interoperabilità nel mondo Linked Data, la possibilità per un esperto del dominio di definire una rappresentazione di alto livello che astrae la conoscenza di dominio dal mezzo tecnico che la gestisce.

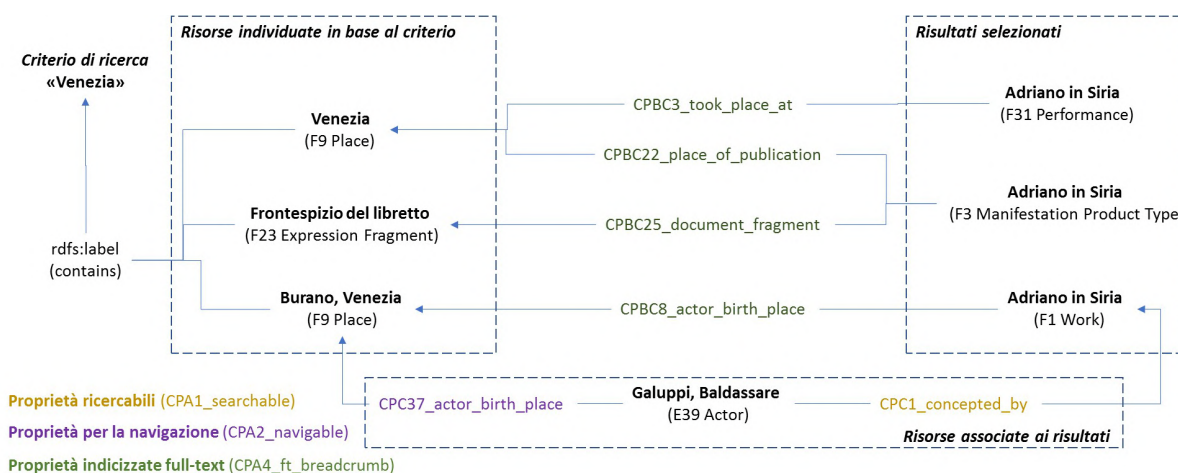


Figura 3

La figura 3 riporta un esempio di ricerca di contenuti testuali. Ciascuna risorsa viene indicizzata in base ad attributi definiti da una specifica tipologia di proprietà (*CPA4\_ft\_breadcrumb*). In questo modo è possibile controllare in base a quali criteri testuali ciascuna tipologia di entità sarà individuabile attraverso la ricerca full-text. Per la ricerca puntuale in base ad attributi specifici, il modello permette all'utente di selezionarli scegliendo tra quelli definiti attraverso una classe di proprietà dedicata (*CPA1\_searchable*). In fase di rappresentazione della risorsa individuata, un'ulteriore classe specifica di proprietà (*CPA2\_navigable*), identifica quali siano gli attributi del grafo rilevanti per comporre la scheda descrittiva che il sistema presenta all'utente.

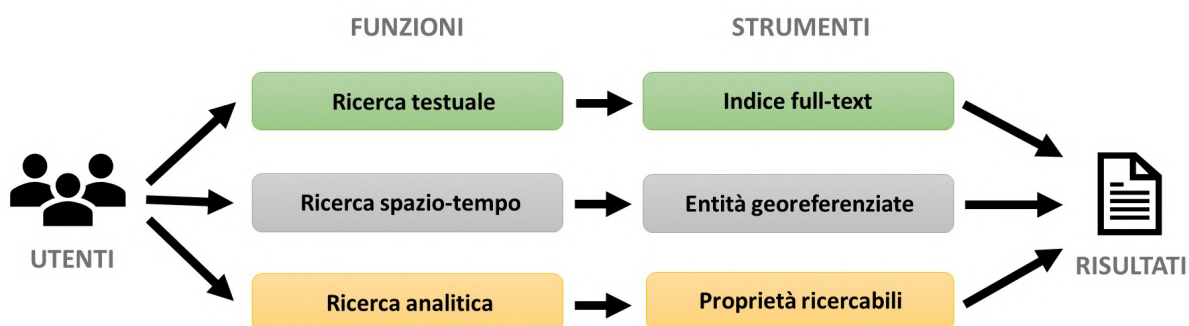


Figura 4

navigazione. Complessivamente lo strato di astrazione semantica è costituito da oltre 130 proprietà suddivise tra proprietà per la navigazione, per la ricerca e l'indicizzazione full-text.

L'accesso ai contenuti è mediato da funzioni di ricerca che rispondono a tre diverse tipologie di criteri: ricerca di un testo libero, collocazione spazio-temporale o caratteristiche specifiche della risorsa ricercata (figura 4). Gli strumenti utilizzati dalle tre modalità sono parte del modello generale definito nel Corago SM: diventa per ciò possibile integrare i diversi criteri per effettuare ricerche che ne contemplino più di uno. Ad esempio, può essere associato il criterio geografico assieme ad un criterio testuale per cercare in quali località del nord Europa è stato rappresentato l'Otello di Verdi. L'esempio evidenzia come questo approccio permetta di integrare contenuti provenienti da altre basi di conoscenza, in questo caso i dati di localizzazione di luoghi e teatri, per estendere le funzionalità proposte dal sistema ai propri utenti. L'estensione del perimetro informativo iniziale attraverso l'allineamento con fonti informative di settore, come il RISM<sup>5</sup>, è un altro caso in cui è possibile utilizzare procedure non supervisionate. In questa prospettiva, l'adozione di ontologie di riferimento estremamente analitiche come FRBRoo e CIDOC CRM, rende più facilmente percorribile una interoperabilità diretta tra basi di conoscenza settoriali [28][29].

L'utilizzo degli strumenti del Web Semantico e dei Linked Data [5] sembra rendere concretamente percorribile la prospettiva di una catalogazione distribuita. La possibilità di verificare la praticabilità di questa prospettiva è, ad oggi, limitata dalla mancanza di dataset necessari a coprire per intero una attività di catalogazione in ambito musicale. Ad esempio, se per gli autori è già disponibile il VIAF<sup>6</sup>, per il repertorio musicale manca ancora un'authority file equivalente. L'adozione di questi strumenti e delle relative metodologie all'interno dei sistemi informativi per i beni culturali potrà favorire la costituzione dell'ecosistema informativo necessario alla realizzazione della catalogazione partecipata e distribuita. Nella necessaria sperimentazione dell'impiego dei Linked Data nell'attività catalogografica sarà determinante il ruolo guida che gli enti nazionali ed internazionali di riferimento dovranno assumere.

## BIBLIOGRAFIA

- [1] Berners-Lee T., Hendler J., Lassila O. 2001. *The Semantic Web*. Scientific American.
- [2] Bartolini I., Luzzi C., Patella M., Pompilio A. e Ruini C. *MIMESIS: A Semantic-based Approach for a Digital Library of Poetry in Music*, in *Proceedings ECLAP 2012: Conference on Information Technologies for Performing Arts, Media Access and Entertainment: Florence 7-9 may 2012* – Firenze, Firenze University Press, 2012, ISBN 9788866551256
- [3] Bellini P. e Nesi P. 2014. *Modeling Performing Arts Metadata and Relationships in Content Service for Institutions*. Multimedia Systems. Springer. DOI:10.1007/s00530-014-0366-0
- [4] Bianconi L., Pompilio A., Pagannone G. 2004. *RADAMES: prototipo d'un repertorio e archivio digitale per il melodramma*. Il Saggiatore Musicale. anno XI n.2. Olschki. Firenze
- [5] Bizer C., Heath T., Berners-Lee T. 2009. *Linked Data - The Story So Far*. International Journal on Semantic Web and Information Systems. Special Issue on Linked Data.
- [6] Bonora P., Ossicini C., Raffa G. 2006. *From Relational Metadata Standards to CRM Ontology: a Case Study in Performing Arts*. CIDOC2006 Proceedings.
- [7] Calvanese D., Mosca A., Remesal J., Rezk M. e Rull G. 2015. *A 'Historical Case' of Ontology-Based Data Access*. Proc. of Digital Heritage 2015. IEEE Computer Society Press.
- [8] Cipollone B. 2009. *La catalogazione dei libretti nel sistema di archiviazione digitale Il Corago*. Quaderni Estensi Rivista: Vol. 1.

---

<sup>5</sup> Répertoire international des sources musicales (<http://www.rism.info/home.html>).

<sup>6</sup> VIAF: Virtual Authority File (<http://viaf.org/>).

- [9] Colazzo D., Goasdoué F., Manolescu I., Roatis A. 2014. *RDF Analytics: Lenses over Semantic Graphs*. 23rd International World Wide Web Conference. Apr 2014. Seoul, South Korea.
- [10] Crawford T., Fields B., Lewis D., Page K. 2014. *Explorations in Linked Data practice for early music corpora*. Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries. IEEE Press. ISBN 9781479955695
- [11] Doerr M. e LeBoeuf P. 2007. *Modelling Intellectual Processes: The FRBR-CRM Harmonization*. In: Digital Libraries: Research & Development. First International DELOS Conference. pp. 114-123. Springer Verlag.
- [12] Doerr, M., Bekiari, C., Le Boeuf, P. 2008. *FRBRoo, a Conceptual Model for Performing Arts*. In 2008 Annual Conference of CIDOC. Athens. 15–18 Sep 2008.
- [13] Doerr, M. 2009. *Ontologies for Cultural Heritage*. Handbook on Ontologies. (pp. 463 - 486). Springer Verlag. ISBN 978-3-540-70999-2
- [14] Doerr M., Gradmann S., Le Boeuf P., Aalberg T., Bailly R., Olensky M. 2013. *Final Report on EDM – FRBRoo Application Profile Task Force*. Europeana.
- [15] Doerr, M. 2003. *The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata*. AAI: AI Magazine. Vol. 24. DOI: <http://dx.doi.org/10.1609/aimag.v24i3.1720> .
- [16] Doty C. 2013. *The Difficulty of An Ontology of Live Performance*. InterActions. UCLA Journal of Education and Information Studies. Vol. 9.1.
- [17] Euzenat J. e Shvaiko P. 2007. *Ontology matching*. Springer
- [18] Guerrini M. e Possemato T. 2015. *Linked data per biblioteche, archivi e musei*. Editrice Bibliografica.
- [19] IFLA Study Group. 1998. *Functional Requirements for Bibliographic Records*. IFLA Section on Cataloguing. K. G. Saur München.
- [20] IFLA Study Group. 2008. *Functional Requirements for Authority Data*. IFLA.
- [21] Kobilarov G., Scott T., Raimond Y., Oliver S., Sizemore C., Smethurst M., Bizer C., Lee R. 2009. *Media Meets Semantic Web – How the BBC Uses DBpedia and Linked Data to Make Connections*. In: ESWC 2009. pp. 723–737. Springer.
- [22] Le Boeuf P. 2009. *Ontologies for Cultural Heritage*. In: Handbook on Ontologies. Springer Berlin Heidelberg. ISBN: 978-3-540-70999-2
- [23] Le Boeuf P. 2012. *Towards Performing Arts Information As Linked Data?* In: SIBMAS 2012 Conference: Best Practice Innovative Techniques for Performing Arts Collections, Libraries and Museums. France.
- [24] Lembo D., Rosati R., Ruzzi M. 2014. *Visualization and Management of Mappings in Ontology-based Data Access*. DL 2014 27th International Workshop on Description Logics. CEUR Workshop Proceedings. Vol. 1193.
- [25] Meroño-Peñuela A., Ashkpour A., van Erp M., Mandemakers K., Breure L., Scharnhorst A., Schlobach S., Frank van Harmelen. 2012. *Semantic Technologies for Historical Research: A Survey*. Semantic Web. IOS Press.
- [26] Peroni S., Tomasi F., Vitali F., Zingoni J. 2014. *Semantic Lenses as Exploration Method for Scholarly Articles*. In: Catarci T., Ferro N., Poggi A. (eds) Bridging Between Cultural Heritage Institutions. IRCDL 2013. Communications in Computer and Information Science, vol 385. Springer, Berlin, Heidelberg.
- [27] Pompilio A., Bianconi L., Regazzi F., Bonora P. 2005. *RADAMES: A new management approach to opera: Repertory, archives and related documents*. Proceedings - First International Conference on Automated Production of Cross Media Content for Multi-channel Distribution. DOI: 10.1109/AXMEDIS.2005.42



- [28] Raimond Y. e Sandler M. 2008. *A web of musical information*. ISMIR 2008: Proceedings of the 9th International Conference on Music Information Retrieval. ISBN 9780615248493.
- [29] Weigl D. M., Page K. R., Lewis D., Crawford T., Knopke I. 2015. *Unifiedaccess to media industry and academic datasets: a case study in Early Music*. Extended abstracts for the Late-Breaking Demo Session of the 16th International Society for Music Information Retrieval Conference.

# Le trascrizioni dei sogni dei sopravvissuti di Auschwitz: analisi linguistica e tematica secondo un modello Xml-Tei

Carlo Chiaramonte, Francesco Figoli, Simone Mastrocesare, Daniele Silvi, Daniele  
Tosco, Marco Zanasi

Università di Roma "Tor Vergata", Italia

This study is financed by the Polish Ministry of Science and Higher Education within the frame of the National Programme for the Development of Humanities, in the years 2016-2019

## PRELIMINARI

Nel 1973, un team psichiatrico dell'Accademia medica di Cracovia inviò un questionario ai sopravvissuti di Auschwitz. Le 147 persone che vi hanno risposto hanno fornito il materiale alla base di questo progetto di ricerca: vite quotidiane, angosce e speranze attraverso i racconti dei loro sogni prima, durante e dopo l'imprigionamento. Il *corpus* originale è redatto in lingua polacca.

Essendo venuti in possesso di questo materiale inedito, grazie alla collaborazione con l'Università di Danzica, presentiamo i risultati dell'analisi testuale eseguita sul *corpus* dei racconti dei sogni opportunamente codificati. Abbiamo utilizzato lo standard XML-TEI per descrivere semanticamente e grammaticalmente le relazioni sui sogni ed è stata prodotta una DTD dedicata con appositi marcatori da noi scelti. Il *corpus* originale è stato studiato con un approccio critico (con conseguente apparato) tale come nei testi letterari, e l'analisi è stata perpetrata anche su due attente traduzioni, eseguite dalla nostra equipe, in inglese e in italiano. La ricerca è stata coadiuvata dall'uso del metodo statistico, che raggiunge diversi obiettivi: esplorare semanticamente la struttura del testo (dimensioni, occorrenze, strutture grammaticali prevalenti) ed essere in grado di tornare al testo originale in qualsiasi momento per arricchire le interpretazioni proposte, documentando il percorso analitico nel suo insieme.

Nel nostro studio abbiamo impiegato tecniche di comprensione dei contenuti facendo affidamento alla visione junghiana del sogno come testo prodotto dall'inconscio. Quest'ultimo è stato quindi valutato tenendo presenti i seguenti aspetti:

1. **Composizione** del testo e la definizione dei personaggi in relazione all'autore/narratore;
2. **Organizzazione** temporale del discorso: il narratore (di una storia o di un episodio) sceglie infatti se indicare i fatti seguendo l'ordine in cui si sono verificati o manipolare le sequenze temporali della narrazione (fabula/intreccio);
3. **Organizzazione** emotiva: il testo narrativo non espone una storia in modo obiettivo e lineare ma, in qualche modo, questa è organizzata dal mittente per adattarla al destinatario, utilizzando degli artifici: anacronia, sincronia, ellissi e carica emotiva dei sintagmi, solo per citarne alcuni.

Abbiamo poi valutato la frequenza dei campi semantici, in particolare gruppi di parole (nomi, verbi, aggettivi o avverbi) che sono usati per descrivere una particolare situazione, ambiente o insieme di oggetti, che appartengono alla vita di tutti i giorni e che creano associazioni precise nelle nostre menti.

Infine abbiamo contato le parole che potrebbero essere correlate al Disturbo Post Traumatico da Stress (PTSD) tenendo presente il manuale *DSM V*.

In particolare, il testo è stato valutato secondo le linee generali come segue:

- \_ Luogo, ossia l'ambientazione fisica della scena onirica; è stata fatta una distinzione tra ambiente aperto (P1.1) chiuso (P1.2) o misto (P1.3) o non definito (P2).
- \_ Contesto della narrazione, con particolare attenzione alla sua qualità descrittiva (CON1), emozionale (CON2) o in assenza di contestualizzazione (CON3).
- \_ Riferimento cronologico, ossia la presenza (T1) o meno (T2) di un'osservazione temporale
- \_ Sequenza narrativa, se lineare (SEQ1) o no (SEQ2), ovvero presenza o assenza di flashback
- \_ Tempi verbali, l'uniformità (VT3) o meno dei tempi verbali al passato (VT1) o al presente (VT2).
- \_ Discorso, il racconto del sogno potrebbe contenere una struttura del discorso diretta (SP1), indiretta (SP2) o entrambe (SP3), oppure il sognatore potrebbe descrivere la scena in una posizione al di fuori della sequenza narrativa (SP4).
- \_ Personaggi, il cast dei sogni e la posizione del sognatore nei confronti degli altri personaggi come solo attore (CH1), presenza co-attori (CH2), presenza di personaggi collaterali (CH3), nessun personaggio (P4).
- \_ Situazione, se la scena descritta sia realistica (S1) o fantastica (S2).
- \_ Stato emotivo, se il sogno contenga (E1) o meno (E2) una chiarificazione delle emozioni del sognatore.

## LA CODIFICA E LE TRADUZIONI

Come abbiamo già avuto modo di far notare, il *corpus* originale è stato redatto in lingua polacca e così ci è stato trasmesso, unitamente ad una traduzione in lingua inglese, svolta da una ricercatrice, madrelingua polacca, della stessa Università di Danzica.

Questa traduzione è stata da noi riveduta e corretta con l'ausilio di un esperto italiano di lingua polacca e infine l'intero corpus è stato ancora tradotto in lingua italiana, mediante un confronto tra le due precedenti versioni: polacca ed inglese.

Sostanzialmente ora disponiamo di tre versioni (in, appunto, tre lingue) che sono state tutte codificate in maniera sinottica, facendo corrispondere i fenomeni individuati e marcati e permettendo inoltre un raffronto delle tre versioni.

## CAMPI SEMANTICI

I campi semantici sono stati individuati valutando tutti i lemmi relativi ai cinque sensi (vista, udito, olfatto, tatto e gusto) e misurando la frequenza della loro insorgenza nei sogni. La linguistica cognitiva continua a fornire prove riguardo la stretta relazione tra l'espressione e l'uso del linguaggio e la percezione del mondo fisico; in altre parole, i pensatori cognitivi vedono il linguaggio come un'altra funzione incorporata tra le altre funzioni cognitive e fortemente incarnata.

## L'ELABORAZIONE STATISTICA

L'elaborazione statistica consiste nel confrontare le modalità di alcuni caratteri qualitativi rilevati su due contrapposti campioni: uno di Sopravvissuti (gruppo sperimentale) e l'altro di giovani (gruppo di controllo) e verificare la loro significatività.

I caratteri rilevati riguardano:

1. il "Luogo" suddiviso nelle modalità L1.1, L1.2, L1.3 ed L2.0;
2. il "Contesto" suddiviso nelle modalità CON1, CON2 e CON3;
3. la "Sequenza Narrativa" suddivisa in SEQ1 e SEQ2;

4. il “Riferimento Temporale” concernente CH1 e CH2;
5. i “Personaggi” relativi a P1, P2, P3 e P4;
6. il “Discorso” afferente alle modalità DS1, DS2, DS3 e DS4;
7. il “Tempo Verbale” con T1, T2 e T3;
8. la “Situazione” in S1 e S2
9. la “Emotività” presente con le modalità E1 ed E2.

Per ciascuna modalità dei suddetti caratteri è stata costruita una tabella di frequenze a doppia entrata (2×2) dove le righe riguardano i campioni (sperimentale = S e di controllo = C) e le colonne, la “presenza” e la “assenza” della modalità posta a confronto.

Gli strumenti metodologici adottati riguardano gli Odds Ratios e il c2 le cui proprietà analitiche consentono, come è noto, di effettuare inferenze e, quindi, di valutare la significatività dei risultati ottenuti.

Ricordiamo che gli Odds sono ottenuti dal rapporto tra la probabilità p che l’evento si verifica e la probabilità contraria 1-p che l’evento non si verifica, per cui un Odd Ratio è dato dal rapporto:

$$\text{Odd Ratio} = \frac{\frac{p_s}{1-p_s}}{\frac{p_c}{1-p_c}} = \frac{a_{11} \times a_{22}}{a_{21} \times a_{12}} = RI$$

dove, nel nostro specifico caso: pS e pC rappresentano le frequenze relative (equivalenti alle probabilità) della modalità in studio, rispettivamente, per il campione degli “Sopravvissuti” e per quello dei soggetti di controllo (gruppo di “Controllo”); a11, a22, a21 ed a12 le corrispondenti frequenze assolute riportate nella suddetta tabella, rispettivamente, nella prima riga e prima colonna; nella seconda riga e seconda colonna; nella seconda riga e prima colonna ed, infine, nella prima riga e seconda colonna. Queste ultime frequenze consentono, in virtù della sottostante uguaglianza, di considerare gli Odds Ratios come “rapporti incrociati” (RI).

Da notare che il rapporto incrociato assume l’unità se, e soltanto se, i due campioni risultano tra loro indipendenti rispetto alla modalità considerata.

Haldane J. B. S. ha dimostrato che in questi casi il logaritmo naturale della v.c. Rapporto Incrociato (Ln RI) converge in distribuzione alla curva normale di media  $m = \text{Ln RI}$  e varianza  $s^2 = S^2$ , cioè  $N(\text{Ln RI}, S^2)$ , dove:

$$S^2 = \frac{1}{a_{11}} + \frac{1}{a_{22}} - \frac{1}{a_{21}} - \frac{1}{a_{12}}$$

da cui l’intervallo di confidenza per Ln RI, al livello di significatività 1-a (dove a rappresenta l’errore di I tipo, cioè la probabilità di rifiutare l’ipotesi nulla quando è vera (in questo caso l’ipotesi nulla suppone che “il rapporto incrociato è pari ad uno”), è dato da:

$$\ln RI \pm u_{\alpha} \times \sqrt{S^2}$$

dove  $\pm u_{\alpha}$  rappresenta l’ascissa della curva normale a cui corrisponde la probabilità, pari ad a, della sua coda (test ad una coda), cioè:

$$\Pr \{ \ln RI \pm u_{\alpha} \times \sqrt{S^2} \} = 1 - \alpha$$

ora, passando ai numeri, l’intervallo di confidenza è delimitato dai seguenti due estremi:

$$\exp \{ \ln RI - u_{\alpha} \times \sqrt{S^2} \}; \exp \{ \ln RI + u_{\alpha} \times \sqrt{S^2} \}$$

A questo punto risulta evidente che gli estremi dell’intervallo possono o meno contenere l’unità del rapporto incrociato per cui, se l’unità non è compresa il risultato ottenuto è significativo al livello di significatività 1-a ; viceversa, se l’unità è invece compresa nell’intervallo, allora il risultato ottenuto non è significativo e, conseguentemente, l’esperimento non ha evidenziato la relazione cercata.

Nel nostro specifico caso si è fissato  $a = 0,05$  e, conseguentemente,  $u_{\alpha} = 1,65$ , per cui:  $1-a = 0,95$ .

Per quanto riguarda il  $\chi^2$ , ricordiamo che l'entità di esso è determinata, con riferimento a queste tabelle  $2 \times 2$ , dalla seguente espressione:

$$\chi^2_{1,df} = \frac{N \cdot (a_{11} \times a_{22} - a_{21} \times a_{12})^2}{n_{01} \times n_{02} \times n_{10} \times n_{20}}$$

dove:  $a_{11}$ ,  $a_{22}$ ,  $a_{21}$ ,  $a_{12}$  assumono gli stessi significati sopra descritti;  $n_{01}$ ,  $n_{02}$ ,  $n_{10}$ ,  $n_{20}$  rappresentano le frequenze marginali della tabella, cioè i totali, rispettivamente, della prima colonna, della seconda colonna, della prima riga e della seconda riga; infine  $N$  è il totale generale.

Con il test in questione si vuole verificare l'ipotesi nulla che consiste nel considerare "le frequenze registrate sono indipendenti dai relativi campioni (Schizofrenici e Controlli)" per cui, si rifiuterà questa ipotesi se il  $\chi^2$  calcolato, con un grado di libertà (g.l.=1), cade nella zona di rifiuto predeterminata ad un livello di significatività pari ad  $\alpha$ ; se, viceversa, il  $\chi^2$  calcolato non cade nella zona di rifiuto, allora non si rifiuta l'ipotesi nulla e, quindi, si conclude nel dire che l'esperimento non è stato idoneo ad evidenziare la relazione di dipendenza cercata.

Ciò premesso, possiamo ora applicare i suddetti modelli alle singole modalità dei caratteri sopra evidenziati riportando, per ogni corrispondente tabella, gli elementi di calcolo ed i relativi risultati. Per tutti gli intervalli di confidenza si è prescelto un livello di significatività, come detto sopra, pari al 95%. Seguono le Tabelle di contingenza:

<b>IPOTESI STATISTICHE:</b>																				
<b>H0 : (ipotesi privilegiata) =  OR  =1</b>												<b>TEST COMBINATI (CLOSED TESTING)</b>								
<b>H1 : (ipotesi alternativa) =  OR  &gt; 1</b>												<b>CONTROLLO DELLA MOLTIPLICITA'</b>								
<b>I. - carattere: "Riferimento Temporale"</b>				<b>IPOTESI MINIMALI</b>								<b>METODO DI BONFERRONI</b>								
<b>Campioni</b>	<b>Modalità</b>			<b>T1</b>	<b>Altro</b>	<b>Tot</b>	OR = 0,03	Ln(Var OR) = 0,093	SE_Ln(Var OR) = 0,305	Z = -11,497	P_value = < 0,001	→	COMMENTO	P_VALUE = 0,001	S < C					
	<b>S</b>	20	106																	126
	<b>C</b>	217	34																	251
	<b>Tot</b>	237	140																	377
IC 95%		ESTR_INF	ESTR_SUP																	
		0,017	0,055																	
<b>Campioni</b>	<b>Modalità</b>			<b>T2</b>	<b>Altro</b>	<b>Tot</b>	OR = 33,826	Ln(Var OR) = 0,093	SE_Ln(Var OR) = 0,305	Z = 11,545	P_value = < 0,001	→	COMMENTO	P_VALUE = 0,001	S > C					
	<b>S</b>	106	20																	126
	<b>C</b>	34	217																	251
	<b>Tot</b>	140	237																	377
IC 95%		ESTR_INF	ESTR_SUP																	
		18,605	61,5																	
<b>2. - carattere: "Contesto"</b>																				
<b>Campioni</b>	<b>Modalità</b>			<b>CON1</b>	<b>Altro</b>	<b>Tot</b>	OR = 4,604	Ln(Var OR) = 0,056	SE_Ln(Var OR) = 0,237	Z = 6,443	P_value = < 0,001	→	COMMENTO	P_VALUE = 0,001	S > C					
	<b>S</b>	88	38																	126
	<b>C</b>	84	167																	251
	<b>Tot</b>	172	205																	377
IC 95%		ESTR_INF	ESTR_SUP																	
		2,893	7,326																	
<b>Campioni</b>	<b>Modalità</b>			<b>CON2</b>	<b>Altro</b>	<b>Tot</b>	OR = 0,456	Ln(Var OR) = 0,089	SE_Ln(Var OR) = 0,298	Z = -2,635	P_value = 0,004	→	COMMENTO	P_VALUE = 0,012	S < C					
	<b>S</b>	17	109																	126
	<b>C</b>	64	187																	251
	<b>Tot</b>	81	296																	377
IC 95%		ESTR_INF	ESTR_SUP																	
		0,254	0,818																	
<b>Campioni</b>	<b>Modalità</b>			<b>CON3</b>	<b>Altro</b>	<b>Tot</b>	OR = 0,584	Ln(Var OR) = 0,078	SE_Ln(Var OR) = 0,279	Z = -1,928	P_value = 0,027	→	COMMENTO	P_VALUE = 0,135	n.s					
	<b>S</b>	21	105																	126
	<b>C</b>	64	187																	251
	<b>Tot</b>	85	292																	377
IC 95%		ESTR_INF	ESTR_SUP																	
		0,338	1,009																	
<b>3. - carattere: "Discorso"</b>																				
<b>Campioni</b>	<b>Modalità</b>			<b>DIS1</b>	<b>Altro</b>	<b>Tot</b>	OR = 2,228	Ln(Var OR) = 0,062	SE_Ln(Var OR) = 0,249	Z = 3,217	P_value = < 0,001	→	COMMENTO	P_VALUE = 0,001	S > C					
	<b>S</b>	42	84																	126
	<b>C</b>	46	205																	251
	<b>Tot</b>	88	289																	377
IC 95%		ESTR_INF	ESTR_SUP																	
		1,368	3,63																	
<b>Campioni</b>	<b>Modalità</b>			<b>DIS2</b>	<b>Altro</b>	<b>Tot</b>	OR = 0,091	Ln(Var OR) = 0,365	SE_Ln(Var OR) = 0,604	Z = -3,968	P_value = < 0,001	→	COMMENTO	P_VALUE = 0,001	S < C					
	<b>S</b>	3	123																	126
	<b>C</b>	53	198																	251
	<b>Tot</b>	56	321																	377
IC 95%		ESTR_INF	ESTR_SUP																	
		0,028	0,297																	

Campioni	Modalità			OR = 0,516 Ln(Var OR) = 0,329 SE_Ln(Var OR) = 0,574 Z = -1,153 P_value = 0,125 ESTR_INF ESTR_SUP IC 95% 0,168 1,589	→	COMMENTO	P_VALUE = 0,865 n.s. S < C	
	DIS3	Altro	Tot					
	S	4	122					126
	C	15	236					251
Tot	19	358	377					
Campioni	Modalità			OR = 1,308 Ln(Var OR) = 0,049 SE_Ln(Var OR) = 0,221 Z = 1,215 P_value = 0,013 ESTR_INF ESTR_SUP IC 95% 0,848 2,017	→	COMMENTO	P_VALUE = 0,052 n.s. S > C	
	DIS4	Altro	Tot					
	S	77	49					126
	C	137	114					251
Tot	214	163	377					
4 - carattere: "Emotività"								
Campioni	Modalità			OR = 0,439 Ln(Var OR) = 0,051 SE_Ln(Var OR) = 0,226 Z = -3,643 P_value = < 0,001 ESTR_INF ESTR_SUP IC 95% 0,282 0,684	→	COMMENTO	P_VALUE = 0,001 S < C	
	E1	Altro	Tot					
	S	44	82					126
	C	138	113					251
Tot	182	195	377					
Campioni	Modalità			OR = 2,276 Ln(Var OR) = 0,051 SE_Ln(Var OR) = 0,226 Z = 3,639 P_value = < 0,001 ESTR_INF ESTR_SUP IC 95% 1,461 3,544	→	COMMENTO	P_VALUE = 0,001 S > C	
	E2	Altro	Tot					
	S	82	44					126
	C	113	138					251
Tot	195	183	377					
5 - carattere: "Luogo"								
Campioni	Modalità			OR = 1,298 Ln(Var OR) = 0,071 SE_Ln(Var OR) = 0,266 Z = 0,981 P_value = 0,164 ESTR_INF ESTR_SUP IC 95% 0,771 2,186	→	COMMENTO	P_VALUE = 0,999 n.s. S > C	
	L1.1	Altro	Tot					
	S	29	97					126
	C	47	204					251
Tot	76	301	377					
Campioni	Modalità			OR = 0,362 Ln(Var OR) = 0,065 SE_Ln(Var OR) = 0,255 Z = -3,985 P_value = < 0,001 ESTR_INF ESTR_SUP IC 95% 0,22 0,597	→	COMMENTO	P_VALUE = 0,001 S < C	
	L1.2	Altro	Tot					
	L	26	100					126
	C	105	146					251
Tot	131	246	377					
Campioni	Modalità			OR = 0,572 Ln(Var OR) = 0,107 SE_Ln(Var OR) = 0,327 Z = -1,708 P_value = 0,044 ESTR_INF ESTR_SUP IC 95% 0,301 1,086	→	COMMENTO	P_VALUE = 0,264 n.s. S < C	
	L1.3	Altro	Tot					
	S	14	112					126
	C	45	206					251
Tot	69	282	351					
Campioni	Modalità			OR = 3,014 Ln(Var OR) = 0,056 SE_Ln(Var OR) = 0,237 Z = 4,655 P_value = < 0,001 ESTR_INF ESTR_SUP IC 95% 1,894 4,796	→	COMMENTO	P_VALUE = 0,001 S > C	
	L2.0	Altro	Tot					
	S	57	69					126
	C	54	197					251
Tot	111	266	357					

6. - carattere: "Personaggi"										
Campioni	Modalità				OR = 6,872 Ln(Var OR) = 0,088 SE_Ln(Var OR) = 0,297 Z = 6,49 P_value = < 0,001 ESTR_INF 3,839 ESTR_SUP 12,3	→	COMMENTO	P_VALUE = 0,001 S > C		
	P1	Altro	Tot							
	S	47	79	126						
	C	20	231	251						
	Tot	67	310	377						
Campioni	Modalità				OR = 1,076 Ln(Var OR) = 0,059 SE_Ln(Var OR) = 0,243 Z = 0,301 P_value = 0,382 ESTR_INF 0,668 ESTR_SUP 1,732	→	COMMENTO	P_VALUE = 0,999 n.s. S > C		
	P2	Altro	Tot							
	S	36	90	126						
	C	68	183	251						
	Tot	104	273	377						
Campioni	Modalità				OR = 0,497 Ln(Var OR) = 0,053 SE_Ln(Var OR) = 0,23 Z = -3,04 P_value = < 0,001 ESTR_INF 0,317 ESTR_SUP 0,78	→	COMMENTO	P_VALUE = 0,001 S < C		
	P3	Altro	Tot							
	S	39	87	126						
	C	119	132	251						
	Tot	168	219	377						
Campioni	Modalità				OR = 0,125 Ln(Var OR) = 0,282 SE_Ln(Var OR) = 0,531 Z = -3,916 P_value = < 0,001 ESTR_INF 0,044 ESTR_SUP 0,354	→	COMMENTO	P_VALUE = 0,001 S < C		
	P4	Altro	Tot							
	S	4	122	126						
	C	52	199	251						
	Tot	56	321	377						
7. - carattere: "Situazione"										
Campioni	Modalità				OR = 9,037 Ln(Var OR) = 0,092 SE_Ln(Var OR) = 0,303 Z = 7,265 P_value = < 0,001 ESTR_INF 4,99 ESTR_SUP 16,366	→	COMMENTO	P_VALUE = 0,001 S > C		
	S1	Altro	Tot							
	S	111	15	126						
	C	113	138	251						
	Tot	224	153	377						
Campioni	Modalità				OR = 0,111 Ln(Var OR) = 0,092 SE_Ln(Var OR) = 0,303 Z = -7,255 P_value = < 0,001 ESTR_INF 0,061 ESTR_SUP 0,201	→	COMMENTO	P_VALUE = 0,001 S < C		
	S2	Altro	Tot							
	S	15	111	126						
	C	138	113	251						
	Tot	153	224	377						
8. - carattere: "Sequenza Narrativa"										
Campioni	Modalità				OR = 38,372 Ln(Var OR) = 0,193 SE_Ln(Var OR) = 0,439 Z = 8,308 P_value = < 0,001 ESTR_INF 16,23 ESTR_SUP 90,719	→	COMMENTO	P_VALUE = 0,001 S > C		
	SEQ1	Altro	Tot							
	S	120	6	126						
	C	86	165	251						
	Tot	206	171	377						
Campioni	Modalità				OR = 0,026 Ln(Var OR) = 0,193 SE_Ln(Var OR) = 0,439 Z = -8,314 P_value = < 0,001 ESTR_INF 0,011 ESTR_SUP 0,061	→	COMMENTO	P_VALUE = 0,001 S < C		
	SEQ2	Altro	Tot							
	S	6	120	126						
	C	165	86	251						
	Tot	171	206	377						
9. - carattere: "Tempo Verbale"										
Campioni	Modalità				OR = 2,623 Ln(Var OR) = 0,059 SE_Ln(Var OR) = 0,243 Z = 3,968 P_value = < 0,001 ESTR_INF 1,629 ESTR_SUP 4,223	→	COMMENTO	P_VALUE = 0,001 S > C		
	T1	Altro	Tot							
	S	49	77	126						
	C	49	202	251						
	Tot	98	279	377						
Campioni	Modalità				OR = 1,068 Ln(Var OR) = 0,053 SE_Ln(Var OR) = 0,23 Z = 0,286 P_value = 0,386 ESTR_INF 0,68 ESTR_SUP 1,676	→	COMMENTO	P_VALUE = 0,999 n.s. S > C		
	T2	Altro	Tot							
	S	43	83	126						
	C	82	169	251						
	Tot	125	252	377						
Campioni	Modalità				OR = 0,403 Ln(Var OR) = 0,056 SE_Ln(Var OR) = 0,237 Z = -3,835 P_value = < 0,001 ESTR_INF 0,253 ESTR_SUP 0,641	→	COMMENTO	P_VALUE = 0,001 S < C		
	T3	Altro	Tot							
	S	34	92	126						
	C	120	131	251						
	Tot	154	223	377						



## CONTEGGI PRELIMINARI E PRIMI RISULTATI

Il lavoro allo stato attuale può definirsi giunto a metà. Da questa breve, e inesaustiva tabella, ci si può fare un'idea in proposito:

Statistiche di base (traduzione italiana):	
Tot. Sogni tradotti (ita):	216
Tot. Sogni presenti nel file (marcatura solo strutturale):	150
Tot. Sogni schedati (marcati approfonditamente):	126
Tot. Sognatori:	43
Token:	8842
Type:	2376
Hapax:	1527
Tot. Fenomeni PTSD riscontrati: 145	

Dopo l'elaborazione statistica (descritta sinteticamente nel paragrafo precedente e che verrà discussa diffusamente in sede di presentazione) sono emerse differenze significative tra i gruppi analizzati (confrontati con soggetti sani, detti "controlli").

Tali differenze si sono riscontrate maggiormente (e significativamente) per quanto riguarda: Contesto, Organizzazione Temporale, Organizzazione Emozionale; in particolare nei seguenti parametri misurati:

Nei resoconti dei sopravvissuti di Auschwitz:

- Il luogo è più **non** definito
- L'osservazione temporale è **meno** rappresentata
- Contesto, è **più** descrittivo
- Lo stato emotivo è **meno** chiaro
- Il discorso è più **diretto**
- Personaggi: **più** protagonisti da **sol**i
- Scena: **più realistica**
- Sequenza: **più lineare**

L'analisi dei campi semantici ha stabilito che i lemmi della vista erano significativamente **più frequenti** nei report del gruppo di controllo.

Molto interessante è la schiacciante quantità di lemmi relativi al Disturbo Post Traumatico da Stress e questo dato, anche se non ancora supportato da un'appropriata valutazione statistica (per la quale sarà necessario avere un ulteriore numero di sogni tradotti) esprime certamente un'importante tendenza.

Per quanto riguarda la ridotta presenza di riferimenti temporali di ambientazione cronologica del sogno, questo è probabilmente legato al fatto che agli ex prigionieri è stato specificamente chiesto di ricordare i sogni che si sono verificati in momenti specifici (prima, durante e dopo la detenzione) per i quali non era più necessario chiarire i tempi.

I sopravvissuti impostano la fabula in luoghi più realistici (S1), rispetto ai controlli.

I sopravvissuti tendono a caratterizzare l'impostazione dei sogni in modo descrittivo (CON1). In realtà, il contesto dei loro sogni è relativamente povero per quanto riguarda l'immaginazione e gli aspetti emotivi. Coerentemente, una rappresentazione del loro stato emotivo è significativamente meno frequente (E2).

In realtà i resoconti dei sopravvissuti sembrano più una fredda cronaca dei fatti, come dimostra la prevalenza della sequenza lineare SEQ1, piuttosto che una storia sviluppata emotivamente.

Questi risultati, a nostro avviso, sembrano rappresentare lo sviamento dell'emozione del sognatore: evitare le emozioni può essere definito qui come il desiderio di evitare di sperimentare o esprimere sensazioni fisiche, pensieri, impulsi e comportamenti legati a stati emotivi intensi.

Disturbo da stress post-traumatico (PTSD) e sviamento emotivo sono strettamente correlati. Molte persone affette da PTSD cercano di sfuggire alle loro emozioni.

Un altro dato di grande interesse è la differenza tra i sopravvissuti e i controlli sani nei campi semantici della Vista: i sopravvissuti mostrano una riduzione dei lemmi legati alla Vista rispetto ai controlli sani.

I nostri risultati suggeriscono che la differenza tra sopravvissuti e controlli nel campo visivo semantico può essere correlata a una compromissione della memoria "visuospatiale" piuttosto che a una riduzione delle esperienze visive nei sognatori depressi.

Questo risultato, e l'enorme quantità, come abbiamo notato sopra, di lemmi riferiti al PTSD nei resoconti dei sogni dei sopravvissuti supporta, a nostro avviso, l'ipotesi che nel gruppo di sopravvissuti ci sia una sostanziale incidenza di Disturbi Post Traumatici da Stress con i correlati sintomi di depressione dell'umore.

Abbiamo prodotto delle tabelle che riassumono i risultati dell'analisi statistica sui dati linguistici e tematici codificati secondo gli schemi XML-TEI. Queste non vengono riportate per ragioni di spazio, ma saranno approfonditamente discusse in sede di presentazione.

## BIBLIOGRAFIA

- [1] Bell, A., Hall, C., (1971). *The personality of a child molester: An analysis of dreams*. Chicago: Aldine Books.
- [2] Bolasco S, Pavone P. (2007). Automatic dictionary and rule-based systems for extracting information from text. In: *Classification and Data Analysis. Book of short paper*, Edizioni Università di Macerata EUM, 255-258.
- [3] Brenneis, C.B., (1971). Features of the manifest dream in schizophrenia. *J Nerv Ment Dis.* 153:81-91.
- [4] Chatman S. (1978). *Story and discourse. Narrative structure in fiction and film*, Ithaca and London, Cornell University Press.
- [5] Domhoff, G. W. (2000). The Repetition Principle in Dreams: Is It a Possible Clue to a Function of Dreams? [http://www.dreamresearch.net/Library/domhoff\\_2000b.html](http://www.dreamresearch.net/Library/domhoff_2000b.html).
- [6] Domhoff, G.W. (1996). *Finding meaning in Dreams. A quantitative approach*. New York: Plenum Press
- [7] Genette, G., (1972). *Figures 3*. Paris: Editions du Seuil.
- [8] Gigliozzi, G. (1997). *Il testo e il computer. Manuale di informatica per gli studi letterari*. Milano: Bruno Mondadori.
- [9] Hall, C., Lind, R. (1970). *Dreams, life and literature: A study of Franz Kafka*. Chapel Hill, NC: University of North Carolina Press.
- [10] Hall, C., Nordby, V. (1972). *The individual and his dreams*. New York: New American Library.
- [11] Hinrichs, E. (1986). Temporal anaphora in discourses of english. *Linguistics and Philosophy*, 9 (1), 63-82.
- [12] Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
- [13] Johnson, M. (1991). Knowing through the body. *Philosophical Psychology*, 4(1), 3 - 18.
- [14] Jung, C. G. (1945). *L'essenza dei Sogni (Vol. 8)*. Torino: Boringhieri.

- [15] Kamp, H., & Rohrer, C. (1983). Tense in texts. In R. Bauerle, R. Schwarze & A. von Stechow (Eds.), *Meaning, Use and Interpretation of Language*. Berlin: de Gruyter.
- [16] Kilroe P. (2000). The dream as text, the dream as narrative. *Dreaming*; 10(3): 125-138.
- [17] Kramer, M., Roth, T., Czaya, J., 1975., Dream development within a REM period. In P. Levin and W. Koella (Eds.), *Sleep 1974: The 2nd European congress on sleep research*, Rome, 1974 (pp. 406-08. Basel: Karger.
- [18] Lakoff, G. (1987). *Women, fire and dangerous things. What categories reveal about the mind* (1st ed.). Chicago-London: The University of Chicago Press.
- [19] Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- [20] Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh : the embodied mind and its challenge to Western thought*. New York: Basic Books.
- [21] Langacker, R. W. (2002). *Concept, image, and symbol : the cognitive basis of grammar* (2nd ed.). Berlin ; New York: Mouton de Gruyter.
- [22] Partes, B. (1984). Nominal and temporal anaphora. *Linguistics and Philosophy*, 7(3), 243-286.
- [23] Schredl M, Erlacher D. (2003) The problem of dream content analysis validity as shown by a bizarreness scale. *Sleep and Hypnosis*; 5: 129-135.
- [24] Schredl, M., & Hofmann, F. (2003). Continuity between waking activities and dream activities. *Consciousness and Cognition* Volume 12, Issue 2, June 2003, 298-308
- [25] Segre, C. (1985). *Avviamento all'analisi del testo letterario*. Einaudi: Torino.
- [26] Silvi D. (2018). *Letteratura Transmediale. Teoria e pratica della cultura digitale*. Bordeaux: Roma, 161-261
- [27] Smith, C. S. (2001). Discourse modes: aspectual entities and tense interpretation. *Cahiers de Grammaire*(26), 183-306.
- [28] Toolan M. (1988). *Narrative: a critical linguistic introduction*. London and New York: Routledge
- [29] Wildes, J. E., Ringham, R. M., & Marcus, M. D. (2010). Emotion avoidance in patients with anorexia nervosa: initial test of a functional model. *The International journal of eating disorders*, 43(5), 398-404.
- [30] Zanasi, M., Calisti, F., Di Lorenzo, G., Valerio, G., & Siracusano, A. (2011). Oneiric activity in schizophrenia: Textual analysis of dream reports. *Consciousness and cognition*, 20(2), 337-348.
- [31] Zanasi, M., Chiaramonte, C., Paoletti, G., Testoni, F., & Siracusano, A. (2010). Oneiric activity in Anorexia. *Sleep and Hypnosis*, 12(1-2), 1-12.
- [32] Zanasi, M., De Persis, S., Caporali, M., & Siracusano, A. (2005). Dreams and age. *Perceptual and motor skills*, 100(3 Pt 2), 925-938.
- [33] Zanasi, M., Pecorella, M., Chiaramonte, C., Niolu, C., & Siracusano, A. (2008). Dreams by persons with mood disorders. *Psychological reports*, 103(2), 381-394.

# Constructing Narratives Using NBVT: A Case Study

Valentina Bartalesi, Daniele Metilli, Carlo Meghini

ISTI-CNR, Via Giuseppe Moruzzi 1, 56124 Pisa  
{valentina.bartalesi,daniele.metilli,carlo.meghini}@isti.cnr.it

## ABSTRACT

Narratives are a fundamental part of human life, starting from the epic poems of the ancient past to modern films. Since the 1970s, much research has been carried out to study the computational representation of narratives. Up to now, there is no standard definition of narrative. In our research, we intend narratives as networks of events defined by a narrator, endowed with participating entities (e.g. persons, location, time) and semantic relations. In this paper, we introduce the Narrative Building and Visualising Tool (NBVT), a semi-automatic software based on a formal ontology for narratives we developed. The tool allows users to construct and visualise narratives using Wikidata as reference knowledge base and Europeana for enriching the narrative with digital objects. As case study, we present the narrative of the life of the Austrian painter Gustav Klimt created using NBVT. Since Wikidata is not event-based, our efforts focus on the automatic extraction from Wikidata of the implicit events that compose the narrative. Furthermore, we developed a dedicated functionality in NBVT that finds the Europeana digital objects related to a particular event. This functionality matches the metadata of Europeana digital objects with the event and the participating entities using a similarity algorithm.

## KEYWORDS

Narrative, Semantic Web, Digital Libraries, Wikidata, Europeana

## INTRODUCTION

Narratives are a fundamental part of human life. Every human being encounters countless stories in their everyday life: from the ones told by people in casual conversation, to the epic poems of the ancient past. The reasons for this centrality of narrative are complex and still under study. A widely-held thesis in psychology is that humans make sense of reality by structuring events into narrative [3]. As aptly put by Taylor (1989), “[a] basic condition of making sense of ourselves [is] that we grasp our lives in a narrative” [13].

Despite the importance of narratives in human life, there is currently no way for a machine to understand narrative. Since the 1970s, much research has been carried out to study the computational representation of narratives [9]. However, to date there is still not a standard way to digitally represent narratives, visualise them, or share them on the Web. One issue is the difficulty in defining exactly what a narrative is. According to David Herman, a narrative is “a cognitive structure or way of making sense of experience, as a type of text, and as a resource for communicative interaction” [6]. Other scholars define narrative in a narrower way, e.g. a network of “temporally-indexed representations of events” [11].

In the last few years, we started investigating the introduction of narratives in Digital Libraries (DLs) using Semantic Web technologies. In particular, we developed a formal ontology for representing narratives [2] based on the CIDOC CRM standard vocabulary [4]. On the basis of the ontology, we built a semi-automated Narrative Building and Visualising Tool (NBVT)<sup>1</sup>, allowing users to construct and visualise narratives using Wikidata<sup>2</sup> as

---

<sup>1</sup> <https://dlnarratives.eu/tool.html>

reference knowledge base (KB) [12]. In our study, we intend narratives as networks of events defined by a narrator, endowed with participating entities (e.g. persons, location, time) and semantic relations.

## STATE OF THE ART

Since the 2000s, in the context of DLs, several tools have been developed to organise and visualise digital collections using semantic models. For example, the CultureSampo project [7] developed an application to explore Finnish cultural heritage contents on the Web, based on Semantic Web technologies. The PATHS project [5] created a system that acts as an interactive personalised tour guide through existing digital library collections. The CULTURA project [1] developed a tool to enrich cultural heritage collections with guided paths in the form of short lessons. The CIPHER project [8] built a set of tools to facilitate the development of a narrative structure from existing or new contents.

In order to introduce narratives in DLs, we developed NBVT. In comparison with the existing tools which focus on specific domains of knowledge, NBVT adopts a domain-independent approach. Indeed, NBVT is based upon: (i) an ontology of narratives that aims to guarantee generality, interoperability and reuse; (ii) an open general-purpose knowledge base to populate the ontology model. This approach allows the construction of different types of narratives by different narrators, and it facilitates the fruition of the narratives from the widest possible audience.

## CONSTRUCTING THE NARRATIVE OF THE LIFE OF GUSTAV KLIMT USING NBVT

NBVT was designed for two kinds of narrators: (i) scholars who want to create a narrative starting from a text written by them, or (ii) general narrators (e.g. school teachers, students) who want to create a narrative based on a text written by someone else, or a narrative existing only in the narrator's mind.

In order to provide the narrator with events and entities to populate the narrative, NBVT uses Wikidata as reference KB. Wikidata is an open collaborative KB containing more than 50 million items organized on the basis of an ontology model [14]. Unfortunately, the number of events contained in Wikidata is relatively low, because Wikidata's ontology is not event-based. The knowledge about events is present in Wikidata, but it is generally represented in an implicit way. For instance, the birth of the Austrian painter Gustav Klimt is not represented as an event "Birth of Klimt", but instead the KB contains a statement of the form "Klimt *place of birth* Vienna". To solve this issue, we extracted the events implicitly contained in the KB, analysing all Wikidata properties and compiling a list of the ones that express implicit events. We collected these events in a graph that we call the Wikidata Event Graph (WEG). The graph is used to import events into NBVT.

We also made an experiment to enrich our narratives using the digital objects collected in Europeana<sup>3</sup>. Europeana is the largest European DL, containing descriptions of about 54 million cultural heritage objects in various formats from more than 3,500 European cultural institutions. In particular, as shown in Fig. 1, we developed a dedicated functionality in NBVT that finds the Europeana digital objects related to a particular event. This functionality matches the metadata of Europeana digital objects with the event and the participating entities using a similarity algorithm [10].

---

<sup>2</sup> <https://wikidata.org>

<sup>3</sup> <https://www.europeana.eu>

As case study, we used NBVT to create the narrative of the life of Gustav Klimt. This artist is well-represented in Europeana, where a search for the string “Gustav Klimt” currently returns 353 objects. Since we are not art historians, we decided to build the narrative based on the English Wikipedia page about the painter. For each event of the narrative, NBVT allows reporting entities from Wikidata, digital objects from Europeana, a textual description from Wikipedia and images from Wikimedia Commons.

The Klimt narrative is composed of 54 events. 31 of them are connected with Europeana digital objects, and 18 are linked to more than one digital object. The total number of digital objects in the narrative is 127, i.e. 36% of all Klimt-related objects in Europeana. It should be noted that several Europeana objects are not related to Klimt’s biography, e.g. posters, modern objects inspired by Klimt. More than 70% of the entities used in the narrative were imported from Wikidata.

None of the 54 events that compose the narrative of the life of Gustav Klimt was explicitly present in Wikidata. After generating the WEG, the number of events that could be automatically detected in Wikidata is 34 (63% of the total). The first version of the tool, presented in [10], allowed the creation of the narrative in about 3 person-days (7 hours per day). After adding the functionality to generate the WEG and to automatically extract and suggest to the user the relevant Europeana digital objects, the time was reduced to 10 hours.

The results of the case study show that NBVT constitutes a significant advancement for the task of digital narrative creation. Indeed, the narrator is able to easily import historical events and related entities from Wikidata and the WEG, thereby reducing the narrative creation workload in a significant way. Furthermore, the linking among different knowledge bases is facilitated since the narrator can easily enrich the narrative with any digital objects contained in the largest European digital library, and with text and images from Wikipedia and Wikimedia Commons.

Currently we are working to develop another significant feature for NBVT: the ability to import texts in natural language and extract events from them in an automated way.

## CONCLUSIONS

In the context of the Digital Humanities, and in particular of Digital Libraries, the narration of major cultural or historical events is a very central point. In this article we have presented NBVT, a semi-automatic tool that allows the creation of narratives that can be used in Digital Libraries.

Different kinds of users can benefit from our approach to create narratives. In particular, users could be: (i) scholars, such as historians, who can create and access narratives about the life and works of the authors they study; (ii) professors who want to use NBVT as a learning tool. A professor may create a narrative on a topic of study and show it to the students. At the same time, the tool could be used by the professor to verify the students’ comprehension by asking them to create a narrative on a particular topic; (iii) exhibition or museum curators who can create a narrative that could be used during a monographic exhibition in order to associate the works of an artist to her/his biography and help the visitors to better understand the life and works of the artist.

NBVT is based on a general ontology for narratives and on a general-purpose KB, i.e. Wikidata. Furthermore, it allows enriching the events of the narratives through Europeana digital objects. NBVT is domain-independent, thus it allows creating different types of narratives, from the history of the giant squid<sup>4</sup> to the evolution of climate change<sup>5</sup>.

---

<sup>4</sup> <https://dlnarratives.eu/timeline/squid.html>

<sup>5</sup> <https://dlnarratives.eu/timeline/climate.html>

As future work, we have planned to explore the automatic extraction of events from text. Adding this functionality to NBVT would make the narrative-building process faster and the user's work easier. However, automatic event extraction could also introduce inaccuracies in the narrative. We plan to analyse this problem in a further study.

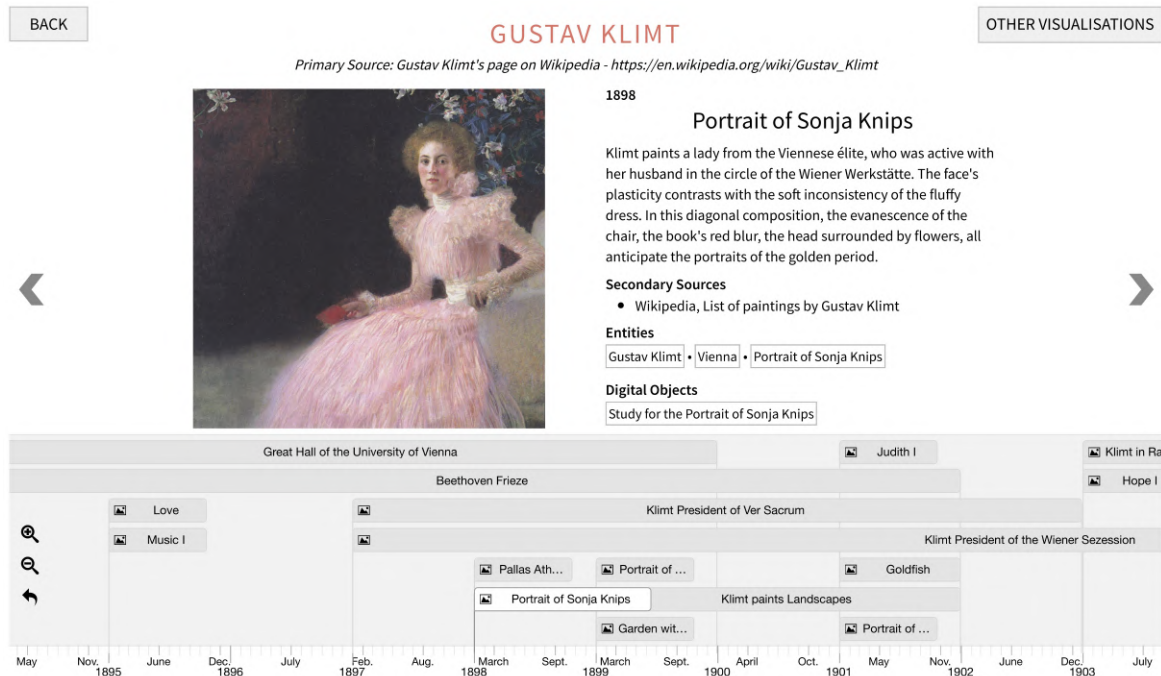


Figure 1. An event in the narrative of Gustav Klimt's life.

## BIBLIOGRAFIA

- [1] Agosti, M., Manoletti, M., Orio, N. and Ponchia, C. 2013. *Enhancing end user access to cultural heritage systems: Tailored narratives and human-centered computing*. In: Proceedings of New Trends in Image Analysis and Processing Conference (ICIAP), pp. 278-287. Springer.
- [2] Bartalesi, V., Meghini, C. and Metilli, D. 2017. *A conceptualisation of narratives and its expression in the CRM*. International Journal of Metadata, Semantics and Ontologies, 12(1), 35–46.
- [3] Bruner, J. 1991. *The narrative construction of reality*. Critical inquiry, 18(1), 1–21.
- [4] Doerr, M. 2003. The CIDOC Conceptual Reference Module: an ontological approach to semantic interoperability of metadata. AI Magazine, 24(3), 75.
- [5] Fernie, K., Griths, J., Archer, P., Chandrinos, K., de Polo, A., Stevenson, M., Clough, P., Goodale, P., Hall, M., Agirre, E., et al. 2012. *PATHS: Personalising access to cultural heritage spaces*. In: 18th International Conference on Virtual Systems and Multimedia (VSMM), pp. 469-474. IEEE.
- [6] Herman, D. 2011. *Basic elements of narrative*. John Wiley & Sons, p. 73.
- [7] Hyvönen, E., Takala, J., Alm, O., Ruotsalo, T. and Mkel, E. 2007. *Semantic Kalevala – accessing cultural contents through semantically annotated stories*. In: Proceedings of the Cultural Heritage on the Semantic Web Workshop at the 6th International Semantic Web Conference (ISWC 2007), Busan, Korea.
- [8] Kilfeather, E., McAuley, J., Corns, A. and McHugh, O. 2003. *An ontological application for archaeological narratives*. In: Proceedings of the 14th International Workshop on Database and Expert Systems Applications, 110-114. IEEE.

- [9] Mani, I. 2014. *Computational narratology*. Handbook of narratology, 84–92.
- [10] Meghini, C., Bartalesi, V., Metilli, D. and Benedetti, F. 2017. *Introducing Narratives in Europeana: Preliminary Steps*. In: *Advances in Databases and Information Systems*, pp. 333-342. Springer, Cham.
- [11] Meister, J. C. 2003. *Computing action: a narratological approach*. De Gruyter, p. xix.
- [12] Metilli, D. 2016. *A Wikidata-based tool for the creation of narratives* (Master's thesis). University of Pisa. (Available from <https://etd.adm.unipi.it>).
- [13] Taylor, C. (1989). *Sources of the self*. Harvard University Press.
- [14] Vrandečić, D. and Krötzsch, M. (2014). *Wikidata: a free collaborative knowledgebase*. *Communications of the ACM*, 57(10), 78–85.



# Progettare, realizzare e promuovere contenuti culturali digitali. Un esempio di didattica integrata: il progetto “Filosofia & Migrazioni”

Maria Eugenia Cadeddu<sup>1</sup>, Francesca Gambetti<sup>2</sup>, Cristina Marras<sup>3</sup>, Ada Russo<sup>4</sup>

<sup>1</sup> CNR-ILIESI, mariaeugenia.cadeddu@cnr.it

<sup>2</sup> Società Filosofica Italiana, Sezione Romana, sfi@sfi.it

<sup>3</sup> CNR-ILIESI, cristina.marras@cnr.it

<sup>4</sup> Ada Russo, CNR-ILIESI, ada.russo@cnr.it

## ABSTRACT

L'intervento presenta il progetto di didattica integrata *Filosofia & Migrazioni. Progettare, realizzare e promuovere contenuti culturali digitali* (<https://filosofiamigrazioni.wordpress.com/>) nato nel 2017 dalla collaborazione tra ricercatrici di storia e di filosofia del Consiglio Nazionale delle Ricerche e della Società Filosofica Italiana - Sezione Romana. Viene descritto il modello di progetto realizzato attraverso il lavoro congiunto di ricercatori, docenti e studenti come esempio di integrazione dei programmi curriculari con lo studio delle migrazioni contemporanee e lo sviluppo di competenze e contenuti digitali. Nella presentazione ci si concentra sulle molteplici connessioni tra mondo della ricerca e mondo della scuola; la complementarità e l'integrazione fra conoscenze e competenze; e l'importanza di un uso critico e consapevole degli strumenti e delle metodologie digitali. Si sottolineano i diversi elementi di criticità individuati soprattutto nell'ancora embrionale processo di alfabetizzazione digitale di docenti e studenti, nell'inadeguatezza delle infrastrutture e della cultura digitale nelle scuole, nella difficoltà ad usare in modo critico le tecnologie.

## PAROLE CHIAVE

Filosofia, Migrazioni, Contenuti culturali digitali, Interdisciplinarietà, Didattica integrata

## IL PROGETTO

Il progetto *Filosofia & Migrazioni. Progettare, realizzare e promuovere contenuti culturali digitali* (<https://filosofiamigrazioni.wordpress.com/>) è nato nel 2017 dalla collaborazione tra ricercatrici di storia e di filosofia del Consiglio Nazionale delle Ricerche e della Società Filosofica Italiana – Sezione Romana. Si è svolto in fase sperimentale nell'ambito delle attività dell'Alternanza Scuola Lavoro 2017-2018 coinvolgendo circa 60 studenti di tre licei: il Liceo delle Scienze Umane “Gelasio Caetani” di Roma, il Liceo Linguistico “Federigo Enriques” di Ostia, e il Liceo Scientifico “Vito Volterra” di Ciampino.

Il progetto ha come pilastro portante la filosofia e il contributo che questa disciplina può dare alla comprensione e all'analisi delle migrazioni intese nel senso più ampio di migrazioni culturali (*translatio studiorum*) (Cadeddu 2015, Cohen 2016, Gregory 2012), e ha sviluppato un modello per lo studio delle migrazioni attraverso l'analisi critico-filosofica di concetti chiave. La terminologia è stata infatti il nucleo di partenza per i percorsi di approfondimento (Cherubini, Faro 2016) attraverso 4 parole chiave: confine, lavoro, conflitto e educazione.

Gli studenti hanno sviluppato una serie di competenze legate alla realizzazione e alla comunicazione di contenuti culturali digitali confrontandosi con temi, testi e problemi relativi alle migrazioni. Sono stati organizzati dei laboratori durante i quali studenti e insegnanti hanno sviluppato competenze informatiche specifiche e acquisito le regole nell'uso pubblico

della comunicazione digitale e la conoscenza di linguaggi e applicazioni, servendosi di dispositivi mobili (tablet, smartphone), oltre ai PC in dotazione alle scuole, e di software free o open source per lo sviluppo dei contenuti (Russo 2018).

Gli insegnanti hanno inserito la riflessione sulle migrazioni nelle loro programmazioni didattiche all'interno di un più ampio progetto di studio e ricerca multidisciplinare (Gambetti, Mastrantonio, Ottaviano 2017). Si è poi cercato di coniugare i contenuti formativi con le realtà territoriali (Cadeddu, Nasso 2016). Gli studenti hanno infatti avuto occasione di visitare “luoghi delle migrazioni” nell'area urbana di Roma, significativi sia per il ruolo culturale svolto, sia per i caratteri di accoglienza: la redazione del giornale PiuCulture, il CAS Salorno, il Centro Caritas di via Venafro a Roma, la Biblioteca Interculturale Cittadini del Mondo.

Per realizzare i contenuti culturali digitali sono state scelte tre diverse strategie di comunicazione ed elaborazione testuale e multimediale: testo non lineare, cronistoria, narrazione. Sono quindi stati identificati i prodotti digitali da realizzare: un ipertesto, in cui il contenuto, prevalentemente di tipo testuale, è stato arricchito da riferimenti ad altri testi o materiali multimediali; una *timeline*, con cui è stata rappresentata graficamente la cronistoria del contenuto da comunicare; un *digital storytelling*, in cui il contenuto da comunicare è stato raccontato attraverso un video e una sequenza di immagini.

Ogni studente ha lavorato ricercando e raccogliendo documentazione su diversi aspetti dell'accoglienza e dell'integrazione culturale, e utilizzando le particolarità dello strumento digitale scelto per valorizzare gli aspetti peculiari delle narrazioni sviluppate. Tutto il lavoro è stato accompagnato dal programma didattico di filosofia attraverso la lettura e lo studio di autori e testi classici del pensiero filosofico e di storia delle idee. Parte integrante dell'intero lavoro è poi stata la realizzazione del sito web di progetto (utilizzando la piattaforma “wordpress” come *content management system*), sviluppato in collaborazione con gli studenti per raccontare l'intera esperienza, pubblicare i contenuti digitali prodotti, documentare le diverse fasi del progetto e i suoi sviluppi, archiviare testi, documenti e materiali.

## IL MODELLO

Come parte integrante del progetto e per strutturare il percorso didattico con un'ottica interdisciplinare, preliminarmente alle attività con gli studenti, si è organizzato un corso di formazione con gli insegnanti nel quale si sono individuati possibili punti di integrazione del programma didattico con l'approfondimento del tema delle migrazioni, nelle sue declinazioni culturali, giuridiche, economiche, ecc. Il modello<sup>1</sup> si presenta come un *workflow* dinamico di lavoro e di organizzazione che integra riflessione teorica con attività pratica, strumenti e materiali, attività esterne con attività d'aula, lezioni frontali, laboratori, e momenti di ‘aula rovesciata’.

L'idea è stata quella di lavorare in modo complesso e articolato, facendo confluire modalità e approcci diversi in relazione alle differenti e successive fasi dello sviluppo del progetto.

Metodologicamente si è scelto di seguire un approccio partecipativo e costruttivistico, basato sulla condivisione e l'accesso aperto a tutti i contenuti e materiali in tutte le fasi del progetto, dalla scelta della piattaforma di condivisione e pubblicazione, ai software e alle applicazioni per l'elaborazione dei contenuti.

Attraverso la realizzazione dei contenuti digitali si è cercato di formare gli studenti ad un uso consapevole e critico delle applicazioni e delle risorse digitali: rigore nei percorsi di ricerca e nella verifica della scientificità e autorevolezza delle fonti, attenzione ai temi della privacy e

---

<sup>1</sup> Una versione semplificata del modello è disponibile tra i materiali dell'Workshop immersivo “Filosofia & Migrazioni”: <https://filosofiamigrazioni.wordpress.com/2018/10/11/un-progetto-di-didattica-integrata-e-il-suo-modello/> (DIDACTA 2019, (<http://fieradidacta.indire.it/>)).

delle regole di uso e diffusione dei contenuti sia testuali che audiovisivi, una guida mirata alla scelta e all'uso dei software e delle applicazioni, e non ultima una riflessione sulla integrazione tra contenuti e percorsi digitali con percorsi di tipo più tradizionale.

L'intera esperienza di questa sperimentazione di un percorso di didattica integrata capace di inglobare anche i contenuti culturali digitali è stata certo positiva sia sul piano dell'apprendimento, sia sul piano dello sviluppo del modello, ma non priva di criticità e di aspetti implementabili, soprattutto per quanto riguarda le modalità di valutazione dei processi e dei risultati del progetto.

## CONCLUSIONI: PROSPETTIVE e CRITICITÀ

Dalla sperimentazione del progetto di didattica integrata sono emersi diversi elementi positivi. Nei questionari finali di valutazione compilati dagli studenti si è evidenziato l'apprezzamento per lo spazio di autonomia che il lavorare con strumenti digitali ha loro offerto, il valore del lavoro di gruppo e l'importanza del confronto con realtà esterne alla scuola. Insegnanti e studenti hanno sottolineato l'impatto positivo e innovativo che il progetto ha avuto nel rinnovamento del loro curriculum scolastico (Gambetti 2018). Quest'ultimo aspetto è ritornato anche nella sintetica griglia di valutazione compilata dai 60 docenti che hanno preso parte al workshop immersivo "Filosofia & Migrazioni" organizzato in occasione di Fiera DIDACTA Italia 2019 (<http://fieradidacta.indire.it/>). Il workshop si è concentrato sull'applicazione del modello del progetto in altri ambiti disciplinari attraverso una serie di esercizi di simulazione. Stante un giudizio positivo sul modello stesso, gli insegnanti hanno sottolineato alcuni elementi interessanti: l'importanza della documentazione e la necessità di materiale più dettagliato e linee guida *ad hoc* per lo svolgimento di progetti didattici integrati a forte componente digitale, il dialogo con il mondo della ricerca (Cadeddu 2015), la scelta critica di metodologie e strumenti digitali in relazione ai contenuti che si intende sviluppare, una adeguata formazione anche degli insegnanti all'uso delle tecnologie.<sup>2</sup>

Nell'intero percorso di progetto la filosofia ha confermato la sua capacità di dialogare e interrogare le categorie e i contesti del nostro presente offrendo strumenti critici e analitici che, sia nella teoria sia nella pratica, permettono di cambiare il modo di fare e insegnare filosofia, anche con l'integrazione di tecnologie digitali (Bynum, Moor 1998, Catalani 2018, Ciraci 2012, Marras 2018, Di Giandomenico 2000).

Tuttavia, accanto ai numerosi aspetti positivi, occorre soprattutto riflettere e analizzare sui diversi elementi di criticità individuati, soprattutto, nell'ancora embrionale processo di mutua alfabetizzazione digitale di docenti e studenti, nell'inadeguatezza delle infrastrutture e della cultura digitale nelle scuole, nella difficoltà ad usare in modo critico e competente le tecnologie.

## BIBLIOGRAFIA

- [1] Bynum, T. W, Moor, J.H. (ed. by). 1998. *The Digital Phoenix. How computers are changing philosophy*, Oxford university Press, Oxford.
- [2] Cadeddu, M. E. 2015. *Il CNR & la scuola. Progetti, ricerche, esperienze*, CNR Edizioni, Roma.
- [3] Cadeddu M. E., Nasso, S. (cura di). 2016. *Reception and integration: the role of small- and medium-sized towns*. CNR Edizioni, Roma.

---

<sup>2</sup> L'elaborazione completa dei dati e le relative tabelle riguardanti i questionari di valutazione sia dei ragazzi sia degli insegnanti che hanno partecipato all'workshop immersivo organizzato per DIDACTA2018 saranno resi disponibili sul sito del progetto.

- [4] Catalani, L. (2018). Informatica e filosofia: contatto!, *Medium*, 23/2/2018, <https://medium.com/@luigicatalani/informatica-e-filosofia-contatto-5dcc2d813cc2>.
- [5] Cherubini M., Faro S., Rinaldi M. (a cura di). 2016. *Glossario sull'asilo e la migrazione*. CNR Edizioni, Roma.
- [6] Cohen S. 2016. *Translatio Studiorum*, In M. Sgarbi (eds) *Encyclopedia of Renaissance Philosophy*. Springer, Cham. [https://doi.org/10.1007/978-319-02848-4\\_1134-1](https://doi.org/10.1007/978-319-02848-4_1134-1)
- [7] Di Giandomenico, M. (a cura di). 2000. *L'uomo e la macchina trent'anni dopo. Filosofia e informatica ieri ed oggi*. Atti del Convegno nazionale della Società Filosofica Italiana, Bari, 24-26 ottobre 1997, Bari.
- [8] Fabio C. 2012. *Informatica per le scienze umane*, McGraw-Hill, Milano.
- [9] Gambetti, F. 2018. Philosophy and Migrations. Designing, implementing and promoting digital cultural contents. Atti della conferenza "Multi-conference on Responsible Research and Innovation in Science", Innovation and Society 2018, September 17-19, Tartu, Springer (forthcoming).
- [10] Gambetti F., Mastrantonio P., Ottaviano G. (a cura di). 2017. *Filosofia e migrazioni. Responsabilità della filosofia e sfide globali*. Diogene Multimedia, Bologna.
- [11] Gregory, Tullio. 2012. *Translatio Studiorum*. In M. Sgarbi (ed.). *Translatio Studiorum, Ancient, Medieval and Modern Bearers of intellectual history*. Brill, Leiden. pp. 1-21.
- [12] Marras, C. 2018. Filosofia, ricerca e innovazione. Considerazioni a margine del documento "Orientamenti per l'apprendimento della filosofia nella società della conoscenza". In *Comunicazione Filosofica*, n. 40, maggio 2018.
- [13] Russo, A. 2018. *Progettare, realizzare e pubblicare contenuti culturali digitali. Vademecum* (<https://filosofiamigrazioni.files.wordpress.com/2018/10/vademecum2.pdf>)

Ultimo accesso agli URL: 3 gennaio 2019

# **Serious Games as immersive educational tools: a model for the Pathways of Romanticism (Porto, Portugal).**

João Victor Camara<sup>1</sup>, Alice Lucas Semedo<sup>2</sup>

<sup>1</sup> Faculdade de Letras da Universidade do Porto – FLUP, Portugal – jvcamara92@gmail.com

<sup>2</sup> Faculdade de Letras da Universidade do Porto – FLUP, Portugal – semedo.alice@gmail.com

## **ABSTRACT**

Drawing on the concept of Serious Games for Cultural Heritage, this paper presents a proposal for a game for the Pathways of Romanticism (Porto, Portugal), a cultural landscape that houses a collection of heritage assets dating from the 19th century. This game aims at providing a creative heritage mediation strategy contributing its cultural dynamization of the Romantic Pathways and promoting a fun and educational immersion experience for its visitors. Serious Games are explored as immersive educational tools and it is argued they have the potential to combine educational and entertainment components while enhancing the visitors' experience on cultural spaces. In addition, the paper seeks to establish how these games can be used both in the context of heritage mediation and humanities education, emphasizing its role in the affective domain of the learning processes involved. Moreover, it establishes a relationship between serious games for cultural heritage and socio-constructivist learning theory, including free-choice learning and situated learning. As a result, the article advances a hypothetical location-based game model whose primary objectives are related to historical processes representation and cultural heritage awareness.

## **KEYWORDS**

Serious Games, Cultural Heritage, Humanities Teaching, Museum and heritage Education, Digital Technologies

## **INTRODUCTION**

In recent years there has been a growing interest in Serious Games and in the potentialities related to its application in different contexts, particularly in heritage sites and museums. As such, this paper dialogues mainly with recent academic production that explore the relationship between Serious Games and Cultural Heritage. Our aim is to provide information about the planning and design of a serious game model for the Pathways of Romanticism (Porto, Portugal). Furthermore, the development of this digital game takes place in the context of a deep tourist and cultural requalification of this space carried out by the Municipality of Porto. The original requalification project foresees the creation of an Interpretation Center, which, in addition to a set of physical interventions in its pedestrian paths, will house a permanent exhibition center and an education unit. The proposal for the design of this game is the result of an investigation carried out during the MA in History and Heritage of the Faculty of Letters of the University of Porto. This research sought to reflect on the implementation of digital games within the cultural sector, more specifically in the context of heritage mediation and humanities education teaching. Thus, the intent of this paper is more to argue for its interest for the heritage mediation of this space, than present its final structure. At this point the development of the game is still at an early stage, its main features are set, and digital technology support means for its operation identified.

## **SERIOUS GAMES AND CULTURAL LANDSCAPES**

Serious Games are digital games that combine instructional and entertainment features in order to provide an immersive experience for players. Zyda (2005) and Anderson et al. (2010) stress that serious games can be implemented in a variety of contexts and disciplines, and they may have primary goals ranging from military training to classroom teaching. Hence, any educational game developed and delivered on a digital platform can be understood as a Serious Game. This research explores Serious Games as a heritage mediation tool that have the potential to enhance the experience and learning of visitors in cultural spaces. Therefore, it suggests the development of an educational mobile game with two main objectives: the representation of historical processes and cultural heritage awareness.

The idea of developing a mobile educational game for a cultural landscape arose from the exploration of the literature on Serious Games for cultural heritage. In general, the authors who investigate the implementation of games in the context of heritage mediation point out a significant set of potentialities. According to Mortara et al. (2014, p. 324), games have the ability to make visits to cultural institutions and sites more dynamic and attractive, increasing the audience involvement with the cultural goods on display. The authors argue that games have a greater impact on the affective domain of the learning process, whereas the feeling of empathy with a character or with the story of the game can favour the understanding of historical processes and different cultures. Avouris and Yiannoutsou (2012, p. 2121) believe the narratives of mobile games can turn a large amount of information from cultural exhibits into an attractive, coherent and didactic activity, as they not only provide information about cultural assets but also enable visitors to organize and control the way the interaction takes place. In the same direction, Bellotti et al (2012, p. 2) argue that games can provide a concrete and convincing context in which players can interact with the cultural objects and contents presented.

## COMMUNICATING CULTURAL HERITAGE WITH DIGITAL TECHNOLOGY



Figure 1: Prototype of a fictional character for the educational game: The water-bearer from Galicia.

Digital technology is a key element of any Serious Game and should be used to enhance the player experience. In the case of educational games developed for cultural landscapes, digital technology also plays an important role in the communication of cultural heritage. For this project, digital technology was considered an appropriate alternative since it allows the communication of various aspects related to the cultural goods of that cultural landscape without the need for significant physical intervention. Besides, the access to mobile networks and the global positioning system allow this communication to occur in real time as visitors explore the site. According to Huizenga, Admiraal, Akkerman and Dam (2009, p. 341), mobile technologies combine situated and active learning with entertainment, and it also increases visitor engagement as it promotes learning in an authentic environment. In addition, Falk & Dierking (2008, pp. 27-28) argue that digital mediation tools provide new multisensory elements that enrich the physical context of experience. Based on these propositions we suggest an educational game for mobile devices (smartphones and tablets) based on the premise of mapping. That is, a game in which the narrative progresses and the interactions take place according to the physical exploration of the cultural landscape by visitors. We believe that the combination of multisensory elements and a narrative based on on-site exploration has the potential to make the communication of cultural heritage more dynamic and attractive.

But what will be communicated in the game and how will digital technology support the communication of cultural heritage? To answer these questions, it is necessary to recall the assertion defended by Mortara et al (2014, p. 324) which establishes that empathy with fictional characters favours the understanding of historical processes and different cultures. Given that the Pathways of Romanticism houses a collection of heritage assets dating from

the 19th century, we selected three historical phenomena that occurred in the city of Porto in this period to be portrayed in the narrative of the game: demographic growth and transformation of the city, the rise of the bourgeoisie as the dominant social group and the transformation of the economic framework with the promotion of the industrial sector. Based on this selection, a review of the academic literature was carried out with the purpose of gathering relevant information about these phenomena to integrate the educational component of the game. Then we selected twelve places of interest located in the cultural landscape. The plan is that interactions between players and characters occur whenever visitors cross these locations. For this reason, they were labeled “crossing points”. Finally, we proposed the creation of twelve fictional characters (figure 1) depicting some of the inhabitants of the Nineteenth-century Porto. Each character will be associated with a specific “crossing point” and will support the communication of aspects related to one of the three historical phenomena portrayed in the narrative of the game.

The interaction between fictional characters and players is the main strategy of communication and interpretation of cultural heritage. The characters will be digitally portrayed on the screen of mobile devices along with other visual pedagogical resources related to the contents and concepts addressed. Audio features also play a central role, since the interaction between characters and players is based on storytelling. Thus, characters share their personal stories every time a visitor walks through one of the “crossing points”. We believe that the multimedia resources employed to tell the stories of the characters can enhance the fantasy aspects of the game and stimulate the curiosity of the visitors. These features can also assist in creating an individual personality for each character, since each one of them will have a unique story, appearance, voice and intonation. This individuality can be used to foster players' empathy with the characters, making them more likely to listen to their stories and consequently pay more attention to the information about the historical phenomena. Moreover, Malone & Lepper (1987, p. 230) stipulate that fantasy - along with challenge, curiosity, and control - is one of the main factors that make learning activities intrinsically motivating. The goal of the players is to discover the largest number of characters and get to know their stories. To find them out, players must wander through the cultural landscape waiting for an invitation from the characters. Note that players do not know how many crossing points exist and where they are located. The only information previously available is the start and end points of the course (Figure 2). Our purpose is for players to feel as if they are embarking on a journey of discovery. In addition, the absence of a pre-established route is intended to grant more freedom and control to players during their experience. Thus, visitors can decide the duration and length of their route according to their motivations, expectations and possibilities.



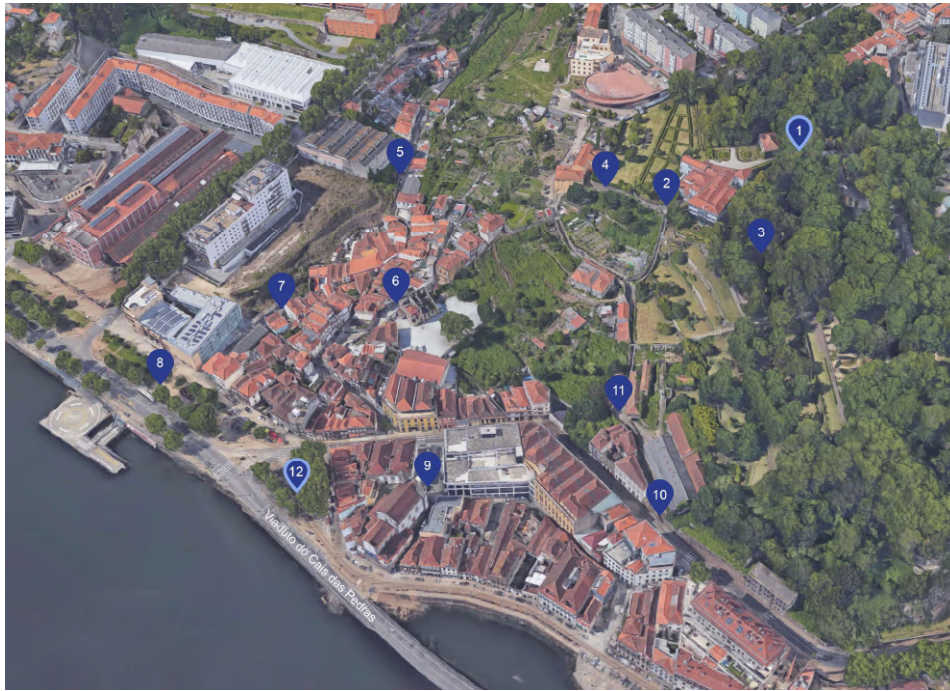


Figure 2: Aerial view of the Pathways of Romanticism with the disposition of the 12 crossing points of the game.

## WHAT COMES NEXT

This paper presents some key features of a serious game for the Pathways of the Romanticism (Porto, Portugal). Over the last year, theoretical research was carried out with the purpose of advancing a viable model of this digital game. Its project development and final design will continue in 2019-20. Among the steps to be taken, the establishment of partnerships with other faculties of the University of Porto aiming at setting up a multidisciplinary team and raising funds for the development of a prototype of the game is seen as an essential one. An application for a PhD scholarship is also being taken under consideration in order to follow up on the development of the game. Having said that, the main goal for the following stage is to bring together professionals and researchers who can make the project feasible as well as to design a collaborative evaluation approach with users.

## BIBLIOGRAPHY

- [1] Anderson, E. F, McLoughlin, L., Liarakapis, F., Peters, C., Petridis, P., Freitas, S. (2010). Developing Serious Games for Cultural Heritage: A State-of-the-Art Review. *Virtual Reality*. 14(4), 255-275.
- [2] Avouris, N. & Yiannoutsou, N. (2012). A Review of Mobile Location-based Games for Learning across Physical and Virtual Spaces. *Journal of Universal Computer Science*. 18, (15), 2120-2142.
- [3] Bellotti, F., Berta, R., De Gloria, A., D'Urso, A., e V. Fiore. (2012). A serious game model for cultural heritage. *ACM Journal on Computing and Cultural Heritage*. 5, 4.
- [4] Falk, J. H. & Dierking, L. D. (2008). Enhancing Visitor Interaction and Learning with Mobile Technologies. In: L. Tallon, L. & K. Walker. *Digital Technologies and the Museum Experience: Handheld Guides and Other Media* (19-33). Forbes Boulevard: AltaMira Press.

- [5] Huizenga, J.; Admiraal, W.; Akkerman, S.; Dam, G. (2009). Mobile game-based learning in secondary education: engagement, motivation and learning in a mobile city game. *Journal of Computer Assisted Learning*. 25, 332–344.
- [6] Malone, T. & Lepper (1987). Making Learning Fun: A Taxonomy of Intrinsic Motivations for Learning. In Snow, R. & Farr, M. J. (Eds.). *Aptitude, Learning, and Instruction Volume 3: Conative and Affective Process Analyses*. New Jersey: Hillsdale.
- [7] Mortara, M., Catalano, C. E., Bellotti, F., Fiucci, G., Houry- Panchetti, M. E Petridis, P. (2014) Learning cultural heritage by serious games. *Journal of Cultural Heritage*. 15, 3, 318-325.
- [8] Zyda, M. (2005). From Visual Simulation to Virtual Reality to Games. *Computer*. 38. 9, 25-32.

# A Corpus Linguistic Approach to Pronouns in a Work of Fiction: The Narrative Voices in Conrad's *The Shadow Line*

Giuseppina Balossi

Independent scholar - teacher of English Language - Liceo Scientifico e Musicale 'G.B. Grassi', Lecco, Italia - g.balossi@libero.it

## ABSTRACT

The advent of the digital age has had an enormous impact on the way we research, teach and think about language (Leech 1992: 1). In Digital Humanities, the application of computer-assisted methods can facilitate enquiry of large corpora, lead to discoveries hardly seen with the naked eye and help put interpretation by intuition to test. Such methods can also assist in language investigation of single texts, and make our close reading more effective. This paper<sup>1</sup> aims to suggest how through a quantitative and qualitative intra-textual analysis we may investigate the narrative voices in a work of fiction. The case study under analysis is Conrad's *Bildungsroman The Shadow-Line, A Confession* (1917). In this work, the neat divide between the *I* protagonist-narrator's internal world and the external world lends itself to investigation of the first person pronouns and the other personal pronouns in the whole work. The statistical output to the present analysis is drawn from Wmatrix3, a [web-based environment containing corpus annotation tools for grammatical and semantic analysis](#). The statistical results and their qualitative analysis will contribute to place the *I*-voice at the centre of narration and to identify foregrounded pronoun-patterns between the *I*-voice and the other narrative voices.

## KEYWORDS

Fiction, Conrad, corpus linguistics, intra-textual, quantitative analysis, qualitative analysis.

## INTRODUCTION

In Digital Humanities (Schreibman *et al.*: 2016), corpus-aided methods have been accepted now as a research paradigm, though they are still not so largely employed in pedagogical contexts (Römer: 2011: 205-225) despite the wealth of literature available on the theory and practicability of such methods (see, for example, Giovanelli *et al.*: 2015; Mahlberg and Stockwell: 2016; Burke *et al.*: 2016).

Originally, corpus-assisted methods involved the analysis of a corpus which Biber (2011: 15) defined as "a large collection of principled texts stored on computer [...] designed to represent a textual domain in a language." However, Biber did not exclude the applicability of corpus linguistics also to literary texts, thus infringing the basic assumption of the corpus approach concerning size and representativeness (Biber *et al.* 1998). In this regard, the distinction made by Adolphs (2006: 65-69; see also McIntyre: 2012) between intra-textual and extra-textual analyses may be helpful to distinguish between our focus of investigation on a 'corpus'. If the analysis is carried out on an individual target text, this is intra-textual; while if a target text is compared against a reference corpus of the same author or with a larger corpus or corpora, our analysis is extra-textual.

---

<sup>1</sup> This paper builds upon the paper I presented at the Annual Conference of the Poetics and Linguistics Association (PALA 2015) and published in Pala Proceedings online (Balossi 2015).

Corpus-aided methods can be focused either on function words (*e.g.* articles, auxiliaries, pronouns, *etc.*) or content words (*e.g.* nouns, verbs, adjectives, *etc.*), or on both. Function words have traditionally been object of quantitative studies carried out in the field of authorship attribution (Burrows 1987; Love 2002; Kestemont 2014: 59-66); however they have also been object of quantitative investigations into style in fictional works (Gibbons and Macrae: 2018). Often, quantitative analyses (*e.g.* word frequencies) applied to literary texts are combined with qualitative analyses (*e.g.* concordances). If with frequencies, we can statistically measure the constituents of language, in running concordances we can make qualitative interpretation of quantitative patterns (Biber *et al.*: 1998; Sinclair: 1991; Mahlberg *et al.* 2016: 433-463). In this paper, both procedures are employed to showcase how the narrative voices in a work of fiction, here Conrad's *The Shadow Line* (2017) can be studied through frequencies and concordances.

The statistical output of the personal pronouns is drawn from Wmatrix4 (Rayson: 2009),<sup>2</sup> a web-based environment containing corpus annotation tools for grammatical and semantic analysis. The tool employed in the present analysis is the POS (Part-of-speech) tool, which is a system designed to attribute grammatical tags to the words in a text. The POS automatic analysis can generate word frequency profiles (*i.e.* list of words, their frequency and relative frequency) and POS frequency profiles (*i.e.* the frequency and relative frequency of part-of-speech tags). Wmatrix can also show concordances for each searched item and carry out comparisons between a target corpus and a reference corpus. The POS tagging is performed by the CLAWS tagger (Constituent Likelihood Automatic Word-tagging System; Garside 1987), a system designed to attribute grammatical tags to the word in a text. Semantic annotation is performed by the USAS tagger (UCREL Semantic Analysis System; <http://ucrel.lancs.ac.uk/wmatrix/>), a system designed to assign semantic tags to the words in a text. In order to obtain an automatic analysis in Wmatrix, users have to upload their texts in plain text (TXT),<sup>3</sup> then the Wmatrix Tag Wizard automatically carries out the POS and USAS tagging.

From the statistics obtained of the personal pronouns and their detailed analysis it will emerge that the frequency of the *I* words referring to the narrator-protagonist, is quantitatively higher compared to those of the other pronouns, and that the *I* narrator-protagonist is characterized, in his interaction with the external world, by the *I-he* pattern. The patterns of the different pronouns will also reveal a peculiar interaction between the subjective *I* and the adult world.

## 2. THE RESULTS

Below I provide the statistics of the top most frequently occurring personal pronouns (Table 2.1) and of the others (Table 2.2).<sup>4</sup> Both tables show the occurring pronoun word<sup>5</sup> (first column from left) followed by the frequencies of each pronoun and their relative frequencies relative to all the words occurring in *The Shadow-Line*. The total words occurring in the text, the total number of pronouns (*e.g.* *I*, *my*, *me*, *myself* and *mine*) and their total relative frequency (*e.g.* 5.11% for the *I*-words in Table 2.1) are shown at the bottom of the table.

In order to see who/what the pronouns refer to, I looked at the concordances generated by Wmatrix. Special attention was paid to distinguish between the *I* words referring to the *I* protagonist and those referring to the other characters, occurring in their direct speech. For

---

<sup>2</sup> The new version Wmatrix4, which replaced Wmatrix3 on December 2018, is here employed. Introduction to Wmatrix and its applications can be accessed at <http://ucrel.lancs.ac.uk/wmatrix/>.

<sup>3</sup> Wmatrix is also capable of dealing with texts in HTML, SGML or XML format.

<sup>4</sup> Impersonal pronouns forms (*e.g.* *one*) have not been object of analysis.

<sup>5</sup> By *I* words, *he* words, *etc.*, I refer to all the pronominal forms (*e.g.* *I*, *me*, *my*, *mine* and *myself*) and to their grammatical categories (*e.g.* subject, object, *etc.*).

example, the pronoun *I*, though mainly indicating the narrator-protagonist, also refers to (ca. 83 times out of 1,316) other characters the I-protagonist interacts with in the story.

**Table 2.1: The top most frequently occurring pronouns**

<i>I</i> words	Freq.	Relative freq.	<i>He</i> words	Freq.	Relative freq.
I	1,233	3.02	he	774	1.89
my	410	1.00	his	455	1.11
me	378	0.93	him	252	0.62
myself	52	0.13	himself	47	0.12
mine	14	0.03			
<b>Tot. 40,848</b>	<b>2,087</b>	<b>5.11</b>		<b>1,529</b>	<b>3.74</b>

**Table 2.2: The other pronouns**

<i>It</i> words	Freq.	Relative freq.	<i>You</i> words	Freq.	Relative freq.	<i>We</i> words	Freq.	Relative freq.
it	594	1.45	you	264	0.65	we	100	0.24
its	42	0.10	yourself	7	0.02	our	45	0.11
			ye	2	0.00	us	44	0.11
			yours	2	0.00	ours	1	0.00
						ourselves	1	0.00
<b>Tot. 40,848</b>	<b>636</b>	<b>1.56</b>		<b>275</b>	<b>0.67</b>		<b>191</b>	<b>0.47</b>

<i>They</i> words	Freq.	Relative freq.	<i>She</i> words	Freq.	Relative freq.
them	80	0.20	her	96	0.24
they	69	0.17	she	49	0.12
their	35	0.09	herself	4	0.01
themselves	2	0.00			
<b>Tot. 40,848</b>	<b>186</b>	<b>0.46</b>		<b>149</b>	<b>0.36</b>

As shown in Table 2.1, the most frequently occurring personal pronouns are the *I* words (5.11%) followed by the *he* words (3.74%). The occurrences recorded for the remaining personal pronouns (cf. Table 2.2) are much lower; among these, the most frequently occurring pronoun is the neutral *it/its* (1.56%) followed by a lower frequency of occurrences of the *you* (0.67%), *we* (0.47%) and *they* words (0.46%); The *she* words hold the least frequency of all (0.36%). The *we*, *you* and *they* words, if looked in their concordances, appear to be mainly employed by the *I* narrator-protagonist to refer to a male or other males who, most of the times, is/are part of the crew, and which for the *we* words also include the protagonist.

Instead, the *she* words point in most of the cases (138 occurrences out of 149), as is the convention in English, to the ship the young captain left, and the ship he is appointed to.

An interesting distinction is flagged up between the different frequencies of occurrence in the male+male patterns. First, the *I-he* pattern prevails over the *I-you*, *I-we* and the *I-they* patterns. When these patterns are looked at in their linguistic context, the *he* words, which the *I*-protagonist refers to, represent for him the adult world and include characters such as Captain Kent, the commander of the narrator's previous ship, the experienced retired navigator Mr. Giles, Hamilton, a non-paying loafer and snob (at the beginning of the story), the haunting figure of the dead former captain whose position the young protagonist takes up, but most of the time to the ship-mates Mr. Burns and Ransome.

The high frequency of occurrence of *I* words in *The Shadow Line* may be associated with its genre, the novel of formation, and its narrative mode (*i.e.* first-person narration; Stape and Simmons, in Conrad 2013: 6). Moreover, the patterns of the different pronouns might also suggest the peculiar interaction of the young captain during a voyage marked by a series of difficulties and challenges. He appears to favour the *I*-Mr. Burns (his antagonist) and the *I*-Ransome (his helper) relationships and tends to exclude the *I-they* and the *I-we* relationships. However, these initial interpretations need to be backed up through detailed analysis that takes into account the *I*-narrative voice in its narrative progression, as exemplified in Section 3.

### 3. The *I*-voice in the Narrative

The passages selected for detailed analysis (see Appendix 1) are taken from the beginning, middle and end of the story, each representing a key phase of the protagonist's maturing process: the protagonist's idealistic, optimistic world-view at the beginning of the story, the obstacles he has to face during his voyage from Bangkok to Singapore, and finally the resolution with the ship and the crew taken safely back to harbour. A summary table of the pronouns for each of the sample can be viewed in Appendix 2. Each table shows the occurring pronouns, their parts of speech (POS),<sup>6</sup> the frequency of each pronoun and the total frequency of each category of pronouns, which was calculated against the total number of words occurring in the sample text. The total words occurring in the text, the total number of pronouns and their total relative frequency are given at the bottom of the table.

In Text 1, the *I* narrator is the sole occupant of the narrative. After the general opening statement "Only the young have such moments," the *I* presents itself as the speaker of the succeeding vague statement: "I don't mean the very young." Yet, in the next sentences the explicit *I* becomes invisible through the use of the impersonal *it* (*e.g.* "It is the privilege of early youth to live in advance of *its* days"), or of *one*, which returns, at the beginning of Paragraphs 2 and 3 or is embedded in sentences (*e.g.* "And the time, too, goes on - till *one* perceives ahead a shadow-line"). In Paragraph 4 the *I* pronoun still occupies a rather 'weak' position within the sentences as it post-modifies "such moments of which *I* have spoken," "Rash moments. *I* mean moments." Again, the speaker seems to hide himself behind intentionally generic observations. Yet the speaker's leading role in the story becomes clear in Paragraph 5 through *me* (1)<sup>7</sup> and *my* (3), which make the story more subjective and personal and help clarify that the protagonist is a mariner. At the beginning of Paragraph 5, *I* moves from a generic and summarizing statement ("It wasn't so bad as that with *me*") to a clearer definition of what "*my* action" had entailed ("*My* action, rash as *it* was, had more the character of divorce"), which is spelled out through "*I* threw up *my* job-chucked *my* berth-left the ship."

---

<sup>6</sup> The POS categories of the personal pronouns are derived from UCREL CLAWS7 tagset, and they can be accessed at <http://ucrel.lancs.ac.uk/claws7tags.html>.

<sup>7</sup> Numbers in round brackets refer to the occurrences of pronouns in the sample texts analysed.

From the generic “my action,” the narrator becomes more specific by clarifying the type of action (“my job”) and next, more concretely, the type of job he left “my berth,” (*i.e.* “the ship”), a decision partially justified by the consideration that “she was a steamship [...]”

Text 2 could be regarded as an illustration of the function of *I-he* pattern that, as pointed out, is the protagonist’s most favoured pattern of interaction in the whole story. Indeed, the *I* is no longer focalizing exclusively on his inner self but is continually moving between *I* (18) and the external world by means of the *he* words (*he* (11), *he* (3), *him* (1)), which deictically point either to Mr. Burns or Ransome. In Paragraph 1, through the *I* words (*I* (3), *me* (2)), we can still acknowledge the pivotal role that the *I*-narrator-protagonist occupies despite the fact that focalization is not on himself but rather on his preoccupation about the ill Mr. Burns (mainly referred to through *he* words (5)). Paragraph 2 is mainly characterized by the dialogue taking place between the narrator-protagonist and Ransome. The *I* words (4) occur less frequently compared to those of *he* (*he* (7), *his* (2)). The topic of the dialogue is centred on the condition of Mr. Burns who is present in the Paragraph 2 as *he* (4). The remaining occurrences of the *he* words (*he* (3) and *his* (2)) refer to Ransome in both the dialogue and narration. Overall, the *I* (3) protagonist is downplayed by the occurrence of narrative *he* (10) and dialogic *he* (10).

Contrary to Text 1, where the internal world of the protagonist was foregrounded, and contrary to Text 2, where the *I-he* pattern prevailed, Text 3 displays the occurrence of different pronouns which in the context help connote its quality: the tension of the dramatic voyage is loosening, and the crew-members, except for the *I* narrator-protagonist, seem to be overawed by the urge to land and leave behind the nightmare of the voyage. On the one hand, the ending of the ordeal foregrounds again the exclusive and somehow dependent relationship of the protagonist towards Ransome, as emerged in the analysis of Text 2. In fact, in Paragraphs 1-2, the *I* words occurring refer to either the *I*-protagonist (11) or to Ransome (4), while the *you* (5) and *he* (9) words to Ransome only. Here, we can see how the *I* protagonist’s external use of *I* and internal use (in his dialogue with Ransome and in his internal thought) are centred on the anxiety and fear of having to depart from Ransome, the adult who has supported, and guided him into his entry in the adult world (*e.g.* “*You don’t mean to leave the ship!*” I cried out. “*He* gasped and a look of almost savage determination passed over *his* face”). Through the pronouns (*e.g. he, they, our*) occurring in the remaining part of the passage, the protagonist shifts onto the external world, *i.e.* all his crew-members referred to a “*they*”.

## CONCLUSION

The quantitative investigation of the personal pronouns has strengthened evidence for the *I*-voice as a foregrounded linguistic feature of *The Shadow Line*. The computer has helped disclose patterns of occurrences not so easily noticeable with the naked eye. Moreover, if we wanted to ascertain whether the salience of the *I*-voice is a real style marker in this work (Enkvist: 1973), and whether such salience can be attributed to the genre (*Bildungsroman*), to the narrative form (*novella*), or to the narrative style (1st person narration), we would need to compare this text with other corpora or with other Conrad’s works, and use of specific statistical tests (*e.g.* keyness analysis).

## REFERENCES

- [1] Adolphs, S. 2006. *Introducing Electronic Text Analysis: A practical guide for language and literary studies*. Abingdon: Routledge.
- [2] Anthony, L. 2015. *AntConc* (Version 3.5). [Computer Software] Tokyo: Waseda University.

- [3] Balossi, G. 2014. A Corpus Linguistic Approach to Literary Language and Characterization: Virginia Woolf's *The Waves*. Amsterdam: John Benjamins.
- [4] Balossi, G. 2015. A computer-aided approach to *I and the World* in Conrad's *The Shadow Line*. On-line Proceedings of the Annual Conference of the Poetics and Linguistics Association (PALA). [[http://www.pala.ac.uk/Annual\\_conferences/proceedings/2015/Balossi2015.pdf](http://www.pala.ac.uk/Annual_conferences/proceedings/2015/Balossi2015.pdf)]
- [5] Biber, D., Conrad, S. and Reppen, R. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: CUP.
- [6] Burke, M. Fialho, O. and Zyngier, S. (eds). 2016. *Scientific Approaches to Literature in Learning Environments*. Amsterdam: John Benjamins.
- [7] Conrad, J. 2017. *The Shadow-Line*, Electronic Version, [www.gutenberg.org](http://www.gutenberg.org); January 9, 2016 [EBook #451].
- [8] Conrad, J. [1917] (2013). *The Shadow-Line, A Confession. The Cambridge Edition to the Works of Joseph Conrad*, J. H. Stape & A. S. Simmons (eds.). Cambridge: Cambridge University Press.
- [9] Burrows, J. F. 1987. *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*. Oxford: Clarendon Press.
- [10] Enkvist, N. E. 1971. *Linguistic Stylistics*. The Hague: Mouton.
- [11] [Garside, R. 1987. The CLAWS word-tagging system. In \*The Computational Analysis of English: A Corpus-based Approach\*, Roger Garside, Geoffrey Leech and Geoffrey Sampson \(eds.\), 30-41. London: Longman.](#)
- [12] Gibbons, A. and Macrae, A. (eds). 2018. *Pronouns in Literature: Positions and Perspectives in Language*. Basingstoke: Palgrave MacMillan.
- [13] Giovanelli, M. A., Macrae, F. T. and Cushing, I. (eds.). 2015. *English Language and Literature. A/AL Level for AQA student Book*. Cambridge: OUP.
- [14] Hoover, D. 2016. 'Computational Stylistics and Text Analysis.' In: R. J. Lane, R. Siemens, and C. Crompton (eds), with J. Rybicki and M. Eder, *Doing Digital Humanities*, pp. 123-144. New York: Routledge.
- [15] Kestemont, M. 2014. 'Function Words in Authorship Attribution. From Black Magic to Theory?'. Proceedings of the Third Computational Linguistics for Literature Workshop, co-located with EACL 2014 – the 14th Conference of the European Chapter of the Association for Computational Linguistics (27 April 2014, Gothenburg, Sweden), pp. 59-66.
- [16] Leech, G. 1991. The state of the art in corpus linguistics. In *English Corpus Linguistics*, K. Aijmer & B. Altenberg(eds), 8–30. London: Longman.
- [17] Leech, G. 1992. Corpora and theories of linguistic performance. In Jan Svartvik (ed.), *Directions in corpus linguistics*, 105–122. Berlin & New York: Mouton de Gruyter.
- [18] Love, H. 2002. *Attributing Authorship: An Introduction*. Cambridge: CUP.
- [19] Mahlberg, M. and Stockwell, P. 2016. [Point and CLiC: Teaching literature with corpus stylistic tools](#). In: M. Burke *et al.* (eds.) *Scientific Approaches to Literature in Learning Environments*, 251–267. Amsterdam: John Benjamins.
- [20] Mahlberg, M., Stockwell, P., de Joode, J., Smith, C., and O'Donnell, M. B. 2016. [CLiC Dickens: Novel uses of concordances for the integration of corpus stylistics and cognitive poetics](#). *Corpora*, 11(3), 433–463.
- [21] McIntyre, D. 2012. *Corpora and Literature*. In: *Wiley Blackwell Encyclopedia of Applied Linguistics*. Oxford: Wiley Blackwell.
- [22] Rayson, P. 2009. Wmatrix: a web-based corpus processing environment. Computing Department: Lancaster University. <http://ucrel.lancs.ac.uk/wmatrix/>
- [23] Römer, U. 2011. Corpus Research Applications in Second Language Teaching. *Annual Review of Applied Linguistics*, 31, 205-225. doi:10.1017/S0267190511000055



- [24] Schreibman, S., Siemens R. and Unsworth, J. (eds). 2016. *A New Companion to Digital Humanities*, 2<sup>nd</sup> edn. Oxford: Wiley-Blackwell.
- [25] Scott, M. 2016. *WordSmith Tools* (Version 7). Stroud: Lexical Analysis Software.
- [26] Sinclair, J. McH. 1991. *Corpus, Concordance, Collocation*. Oxford: OUP.

## Appendix 1

### Text 1

Only the young have such moments. *I* don't mean the very young. No. The very young have, properly speaking, no moments. *It* is the privilege of early youth to live in advance of *its* days in all the beautiful continuity of hope which knows no pauses and no introspection.

*One* closes behind *one* the little gate of mere boyishness--and enters an enchanted garden. *Its* very shades glow with promise. Every turn of the path has *its* seduction. And *it* isn't because *it* is an undiscovered country. *One* knows well enough that all mankind had streamed that way. *It* is the charm of universal experience from which one expects an uncommon or personal sensation - a bit of *one's* own.

*One* goes on recognizing the landmarks of the predecessors, excited, amused, taking the hard luck and the good luck together--the kicks and the half-pence, as the saying is--the picturesque common lot that holds so many possibilities for the deserving or perhaps for the lucky. Yes. *One* goes on. And the time, too, goes on - till *one* perceives ahead a shadow-line warning *one* that the region of early youth, too, must be left behind.

This is the period of life in which such moments of which *I* have spoken are likely to come. What moments? Why, the moments of boredom, of weariness, of dissatisfaction. Rash moments. *I* mean moments when the still young are inclined to commit rash actions, such as getting married suddenly or else throwing up a job for no reason.

This is not a marriage story. *It* wasn't so bad as that with *me*. *My* action, rash as *it* was, had more the character of divorce - almost of desertion. For no reason on which a sensible person could put a finger *I* threw up *my* job - chucked *my* berth - left the ship of which the worst that could be said was that *she* was a steamship and therefore, perhaps, not entitled to that blind loyalty which... . However, *it's* no use trying to put a gloss on what even at the time *I myself* half suspected to be a caprice.

(2013: 11-12)

### Text 2

*I* avoided giving Mr. Burns any opening for conversation for the next few days. *I* merely used to throw *him* a hasty, cheery word when passing *his* door. *I* believe that if *he* had had the strength *he* would have called out after *me* more than once. But *he* hadn't the strength. Ransome, however, observed to *me* one afternoon that the mate "seemed to be picking up wonderfully."

"Did *he* talk any nonsense to *you* of late?" *I* asked casually. "No, sir." Ransome was startled by the direct question; but, after a pause, *he* added equably: "*He* told *me* this morning, sir, that *he* was sorry *he* had to bury *our* late captain right in the ship's way, as one may say, out of the Gulf." "Isn't this nonsense enough for *you*?" *I* asked, looking confidently at the intelligent, quiet face on which the secret uneasiness in the man's breast had thrown a transparent veil of care. Ransome didn't know. *He* had not given a thought to the matter. And with a faint smile *he* flitted away from *me* on *his* never-ending duties, with *his* usual guarded activity.

Two more days passed. *We* had advanced a little way - a very little way - into the larger space of the Gulf of Siam. Seizing eagerly upon the elation of the first command thrown into *my* lap, by the agency of Captain Giles, *I* had yet an uneasy feeling that such luck as this has got perhaps to be paid for in some way. *I* had held, professionally, a review of *my* chances. *I* was

competent enough for that. At least, *I* thought so. *I* had a general sense of *my* preparedness which only a man pursuing a calling *he* loves can know. That feeling seemed to *me* the most natural thing in the world. As natural as breathing. *I* imagined *I* could not have lived without *it*.

*I* don't know what *I* expected. Perhaps nothing else than that special intensity of existence which is the quintessence of youthful aspirations. Whatever *I* expected *I* did not expect to be beset by hurricanes. *I* knew better than that. In the Gulf of Siam there are no hurricanes. But neither did *I* expect to find *myself* bound hand and foot to the hopeless extent which was revealed to *me* as the days went on.

(2013: 68-69)

### Text 3

"*You* don't mean to leave the ship!" *I* cried out. "*I* do really, sir. *I* want to go and be quiet somewhere. Anywhere. The hospital will do." "But, Ransome," *I* said. "*I* hate the idea of parting with *you*." *I* must go," *he* broke in, "*I* have the right." *He* gasped and a look of almost savage determination passed over *his* face. For an instant *he* was another being. And *I* saw under the worth and the comeliness of the man the humble reality of things. Life was a boon to *him* - this precarious hard life, and *he* was thoroughly alarmed about *himself*.

"Of course *I* shall pay you off if you wish *it*," *I* hastened to say. "Only *I* must ask you to remain on board till this afternoon. *I* can't leave Mr. Burns absolutely by *himself* in the ship for hours." *He* softened at once and assured *me* with a smile and in *his* natural pleasant voice that *he* understood that very well.

When *I* returned on deck everything was ready for the removal of the men. *It* was the last ordeal of that episode which had been maturing and tempering *my* character - though *I* did not know *it*.

*It* was awful. *They* passed under *my* eyes one after another - each of *them* an embodied reproach of the bitterest kind, till *I* felt a sort of revolt wake up in *me*. Poor Frenchy had gone suddenly under. *He* was carried past *me* insensible, *his* comic face horribly flushed and as if swollen, breathing stertorously. *He* looked more like Mr. Punch than ever; a disgracefully intoxicated Mr. Punch.

The austere Gambriel, on the contrary, had improved temporarily. *He* insisted on walking on *his* own feet to the rail - of course with assistance on each side of him. But *he* gave way to a sudden panic at the moment of being swung over the side and began to wail despairingly:

"Don't let *them* drop *me*, sir. Don't let *them* drop *me*, sir!" While *I* kept on shouting to *him* in most soothing accents: "All right, Gambriel. *They* won't! *They* won't!"

*It* was no doubt very ridiculous. The blue-jackets on *our* deck were grinning quietly, while even Ransome himself (much to the fore in lending a hand) had to enlarge *his* wistful smile for a fleeting moment.

*I* left for the shore in the steam pinnace, and on looking back beheld Mr. Burns actually standing up by the taffrail, still in *his* enormous woolly overcoat. The bright sunlight brought out *his* weirdness amazingly. *He* looked like a frightful and elaborate scarecrow set up on the poop of a death-stricken ship, set up to keep the seabirds from the corpses.

(2013:102-103)

## Appendix 2

**Table 3.1: Pronouns words in Sample Text 1**

Pronouns words	POS	Freq.	Relative Freq.
<i>I</i> words			
I	PPIS1	5	
my	APPGE	3	
me	PPIO1	1	
myself	PPX1	1	
Tot.		10	2.82
<i>It</i> words			
It	PPH1	7	
its	APPGE	3	
Tot.		10	2.82
<i>She</i> words			
she	PPHS1	1	
Tot.		1	0.28
<b>Tot. words</b> 354		<b>Tot. Pronouns</b> 21	5.93

**Table 3.2: Pronoun words in Sample Text 2**

Pronoun words	POS	Freq.	Relative Freq.
<i>I</i> words			
I	PPIS1	18	
me	PPIO1	6	
my	APPGE	3	
myself	PPX1	1	
Tot.		28	7.19
<i>He</i> words			
he	PPHS1	11	
his	APPGE	3	
him	PPHO1	1	
Tot.		15	3.85
<i>We</i> words			
we	PPIS2	1	
our	APPGE	1	
Tot.		2	0.51
<i>You</i> words			
you	PPY	2	
Tot.		2	0.51
<i>It</i> words			
it	PPH1	1	
Tot.		1	0.26
Tot. words 389		<b>Tot. Pronouns 48</b>	12.33

**Table 3.3: Pronoun words in Sample Text 3**

Pronoun words	POS	Freq.	Relative Freq.
<i>I</i> words			
I	PPIS1	17	
me	PPIO1	5	
my	APPGE	2	
Tot.		24	5.32
<i>He</i> words			
he	PPHS1	11	
his	APPGE	7	
him	PPHO1	3	
himself	PPX1	3	
Tot.		24	5.32
<i>They</i> words			
they	PPHS2	3	
them	PPHO2	3	
Tot.		6	1.33
<i>You</i> words			
you	PPY	5	
Tot.		5	1.11
<i>We</i> words			
Tot.	APPGE	1	0.22
<i>It</i> words			
it	PPH1	5	
Tot.		5	1.11
<b>Tot. words</b> 451		<b>Tot. Pronouns</b> 65	14.41

# Corpora in classe – il caso della lingua serba per italofoni

Olja Perišić Arsić

Università degli studi di Torino, olja.perisic@unito.it

## KEYWORDS

corpora, lingua serba, DDL, insegnamento, corpus-based

Nel presente lavoro ci occuperemo dell'uso dei corpora linguistici nell'apprendimento della lingua serba come lingua straniera. Le prime ricerche circa l'impiego dei corpora nell'insegnamento risalgono agli anni '90 e vedono nell'approccio DDL – data driven learning (Johns 1991) la possibilità, da parte di chi apprende, di avere accesso diretto ai dati linguistici e di arrivare in modo sempre più autonomo, attraverso il metodo induttivo, alle regole di funzionamento di una lingua. Con l'applicazione pratica di quanto definito in via teorica si presentavano tuttavia realtà didattiche che non consentivano l'uso dei computer in classe. Dall'approccio *hands-on* si è scoperto che si poteva efficacemente passare a quello *hands-off* (Boulton 2012), che consiste nell'uso di concordanze prestampate, scelte dall'insegnante rispetto all'argomento trattato. Un'altra obiezione circa la metodologia corpus-based è che si tratta di un procedimento complesso che richiede un addestramento tecnologico finalizzato all'apprendimento linguistico. L'uso dei computer in tal caso può rappresentare un ostacolo sia per gli studenti sia per gli insegnanti, i quali non sempre sono propensi all'uso delle tecnologie in classe (Yoon e Hirvela 2004) e alla simultanea gestione di nuovi programmi, nuovi software e di un nuovo approccio allo studio (Boulton 2010, 2012). Gli argomenti a favore dell'approccio *hands-off* rispetto a quello *hands-on* sono giustificati nei casi di classi numerose la cui gestione può presentare difficoltà oggettive. Nelle classi meno affollate (costituite in media da dieci persone circa), come sono quelle di serbo/croato, si può organizzare il lavoro in modalità *hands-on*, tenuto conto della possibilità di seguire e assistere singolarmente ogni studente.

Per stabilire una metodologia corpus-based adatta all'apprendimento del serbo come lingua straniera, si può partire dall'unica ricerca finora svolta in questo ambito (Perišić Arsić 2018a), che ha coinvolto 5 studenti del Dipartimento di Lingue e letterature straniere e culture moderne dell'Università di Torino: 3 della terza annualità e 2 della Laurea Magistrale. In conclusione gli studenti hanno manifestato interesse ed entusiasmo per questo tipo di apprendimento, e hanno prodotto risultati concreti che sono stati esposti nella relazione finale. Per ottenere esiti soddisfacenti sarebbe indispensabile frequentare il corso una volta alla settimana (2 ore di lezione) per un intero anno accademico. Nel lavoro in classe vengono utilizzati i corpora per il serbo e il croato presenti su due piattaforme: la prima è stata sviluppata dalla Facoltà di Matematica di Belgrado, <http://www.korpus.matf.bg.ac.rs/prezentacija/korpus.html>; la seconda, Sketch Engine, è un moderno software che contiene i corpora per 93 lingue (<https://www.sketchengine.eu/>). Entrambe le piattaforme sono ad accesso libero e gratuito previa registrazione. Nella prima fase di lavoro gli studenti apprendono le basi per fare una *query* utilizzando i metacaratteri per richieste semplici e CQL – corpus query language, che consentono di interrogare il corpus per ricerche più elaborate. Nella fase successiva, gli studenti ai primi livelli apprendono le basi della morfologia mediante una serie di *query* per lemma che restituiscono tutte le possibili forme di una parola che, in una lingua molto flessiva come il serbo, sono numerose; in seguito osservano i diversi tipi di concordanze, ad es. Agg.+N, Avv.+V (sarà poi lo studente a scoprire la posizione dell'aggettivo rispetto al nome o dell'avverbio rispetto al verbo) ecc. A

livello intermedio è possibile fare ricerche lessicali e semantiche sulle collocazioni di una parola: i risultati offriranno informazioni preziose sulla combinazione delle parole più frequenti, mentre sotto un profilo contrastivo mostreranno le equivalenze funzionali che i dizionari spesso non trattano in modo esauriente. Per quanto riguarda l'apprendimento dell'aspetto verbale, i corpora si mostrano inoltre particolarmente efficaci grazie anche alla ricerca attraverso un indefinito numero di caratteri che precedono il verbo e che permetteranno di trovare tutte le forme perfettive del verbo stesso ottenute con la prefissazione, ad es. [lemma=".\*raditi"].

Gli studenti della laurea Magistrale possono lavorare sulla traduzione sia settoriale, vista ad esempio la presenza di corpora con testi giuridici come quelli dell'Unione Europea su Sketch Engine (Eur-lex – variante croata), sia letteraria, utilizzando i testi letterari e giornalistici del corpus serbo SrpKor2013. Nella didattica della traduzione i corpora sono utili per colmare l'assenza, nell'ambito della lingua serba, di recenti dizionari elettronici basati su dati estrapolati dai corpora e che permettono da un lato di individuare una parola come *functionally complete unit of meaning* (Tognini Bonelli 1996), dall'altro di trovare non solo equivalenze semantiche, ma funzionali, pragmatiche e stilistiche. Inoltre dall'analisi di alcuni dizionari bilingui si evince la mancanza di esempi autentici (Perišić 2018b) che però può essere facilmente colmata dall'uso dei corpora la cui caratteristica principale è data proprio dall'autenticità di testi in grado di facilitare l'apprendimento di una lingua non solo dal punto di vista lessico-grammaticale, ma a livello funzionale e comunicativo. In mancanza di altre ricerche in questo campo, si spera di poter disporre in futuro di più dati per valutare a livello statistico, quantitativo ma anche qualitativo, l'efficacia e i risultati conseguiti seguendo l'approccio DDL.

Oltre a mostrare le potenzialità d'uso dei corpora nell'ambito della lingua serba come LS, cercheremo di rispondere a due domande, ossia: rispetto agli approcci consueti, il DDL presenta dei vantaggi? E in caso affermativo, quali sono? In molte realtà la didattica tradizionale si mostra tuttora efficace, a prescindere dalla tipologia di studente o di insegnante che si può trovare a suo agio seguendo un orientamento di questo tipo. Per tale motivo cercheremo di porre a confronto i due approcci, al di là delle preferenze individuali o del metodo utilizzato (deduttivo vs. induttivo), basato sul tipo di informazione cui lo studente può accedere esclusivamente mediante l'uso dei corpora e che può rendere l'approccio corpus-based più vantaggioso rispetto a quello tradizionale.

Oggi, grazie allo sviluppo dei software che permettono complesse analisi linguistiche, qualitative e quantitative, il DDL, che studia l'uso dei corpora nelle classi di lingua, apre maggiori e più complete prospettive di indagine non solo in relazione alla lingua inglese. In Serbia l'insegnamento delle lingue attualmente non dà molto spazio a questo criterio (Vítaz e Poletanović 2018), per cui non sorprende il fatto che in una disciplina ancora poco sviluppata come quella del serbo quale L2/LS si registri quasi la totale assenza di ricerche corpus-based.

Con il presente lavoro si intende pertanto:

far luce sull'uso dei corpora nell'insegnamento della lingua serba per gli italofoeni;

presentare il lavoro teorico e pratico svolto finora;

tracciare un possibile itinerario per future ricerche che in una prospettiva interdisciplinare possano coinvolgere lessicografia, traduzione, sociolinguistica e altre discipline affini.

## BIBLIOGRAFIA

- [1] Boulton A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning* 60(3): 534-572.
- [2] Boulton A. (2012). *Hands-on / hands-off: Alternative approaches to data-driven learning*. In J. Thomas, A. Boulton (eds.), *Input, Process and Product: Developments in Teaching and Language Corpora*. Brno: Masaryk University Press, 152-168.

- [3] Johns T. (1991). *Should you be persuaded: Two examples of data-driven learning*. In T. Johns, P. King (ed.), *Classroom Concordancing*. English Language Research Journal 4: 1-16,  
[https://lexically.net/wordsmith/corpus\\_linguistics\\_links/Tim%20Johns%20and%20DDL.pdf](https://lexically.net/wordsmith/corpus_linguistics_links/Tim%20Johns%20and%20DDL.pdf)
- [4] Korpus savremenog srpskog jezika (versione SrpKor2013), Copyright Grupa za jezičke tehnologije Univerziteta u Beogradu (<http://www.korpus.matf.bg.ac.rs>), autori: Duško Vitas e Miloš Utvić.
- [5] Perišić Arsić O. (2018a). *Upotreba korpusa u didaktici srpskog jezika kao stranog* (Utilizzo dei corpora nella didattica del serbo come LS). Lavoro presentato al convegno internazionale “Serbo come lingua straniera in teoria e prassi” IV, Belgrado, 26-28 ottobre 2018.
- [6] Perišić Arsić O. (2018b). *Recensione di “Italijansko-srpski, srpsko-italijanski rečnik za osnovnu i srednju školu di Jasmina Stojković e Katarina Zavištin”*. *Ricognizioni* 5(9): 213-217.
- [7] Sketch Engine, <http://www.sketchengine.co.uk>.
- [8] Tognini-Bonelli E. (1996). *Translation Equivalence in a Corpus Linguistics Framework*. *Rivista internazionale di tecnica della traduzione* 2: 28-53.
- [9] Vitaz M., Poletanović M. (2018). *Data Driven Learning – The Serbian case*. Lavoro presentato alla conferenza “Data Driven Learning for Foreign Languages and Clil Classes” presso l’Università degli studi di Torino, 28/09/2018.
- [10] Yoon H., Hirvela A. (2004). *ESL student attitudes toward corpus use*. *L2. Journal of Second Language Writing* 13(4): 257-283.



# Un'indagine sulle competenze digitali di studenti di facoltà umanistiche nativi digitali di ultima generazione

Floriana C. Sciumbata

Università di Udine/Università di Trieste

## ABSTRACT

Nella sua ormai nota pubblicazione del 2001 *Digital natives, digital immigrants*, Prensky afferma che «our students today are all “native speakers” of the digital language of computers, video games and the Internet». Tuttavia, l'esposizione alla tecnologia non si traduce necessariamente in competenza. Questa proposta vuole perciò indagare le conoscenze digitali di studenti universitari nati dopo il 1993 iscritti a corsi in materie umanistiche, in cui le competenze digitali e l'informatica rischiano spesso di passare in secondo piano sia per quanto riguarda l'interesse degli studenti sia nell'offerta formativa. Si tratta di un'indagine conoscitiva per capire come gli studenti percepiscono il proprio livello di competenza digitale e se esiste una discrepanza tra il livello percepito tramite autovalutazione e la conoscenza effettiva. Perciò, gli studenti hanno risposto anonimamente a un sondaggio divulgato via internet che comprende quattro sezioni: la prima con domande conoscitive per tracciare un profilo dei rispondenti; la seconda in cui si chiede loro di collocare la propria competenza digitale su una scala da insufficiente a eccellente; la terza costituita da un'autovalutazione secondo i criteri del Quadro Europeo delle Competenze Digitali. Infine, la quarta parte comprende quindici domande a risposta multipla, pesate tramite punteggi, su aspetti pratici e teorici del lavoro al computer utili in ambito universitario e professionale: per esempio la gestione e la conversione di file, alcuni aspetti della navigazione su Internet (motori di ricerca, sicurezza, ma anche fonti affidabili per le notizie) e così via. I risultati saranno utili per tracciare il profilo delle capacità digitali degli studenti universitari. Inoltre, permetteranno di individuare quali conoscenze digitali, spesso date per assodate da docenti e istituzioni, dovrebbero invece essere potenziate a scuola e all'università per fornire agli studenti strumenti adeguati ad affrontare il loro percorso accademico e, più avanti, quello professionale, ma anche l'uso quotidiano degli strumenti digitali.

## PAROLE CHIAVE

Didattica, competenze digitali, studenti universitari, materie umanistiche, ITC

Nella sua ormai nota pubblicazione del 2001 *Digital natives, digital immigrants*, Mark Prensky afferma che «our students today are all “native speakers” of the digital language of computers, video games and the Internet». Tuttavia, l'esposizione costante alla tecnologia non si traduce necessariamente in competenza: secondo uno studio di Calvani *et al.* (2012), studenti di 14-16 anni riescono a compiere operazioni procedurali al computer, ma ottengono punteggi più bassi nel caso di operazioni più complesse o di domande collegate al funzionamento della tecnologia, all'applicazione dell'informatica e all'uso di operatori logici (Calvani *et al.* 2012: 799). Inoltre, solo la metà del campione è riuscita a valutare correttamente l'affidabilità di informazioni trovate su internet. Anche un'indagine OCSE-PISA del 2015 sulle competenze digitali rileva che gli studenti italiani «do not learn how to plan and execute a search, how to evaluate the usefulness of information, or how to assess the credibility of sources» (OCSE-PISA 2015), e lo stesso documento sottolinea che le scuole non riescono a fornire gli strumenti necessari per potenziare queste conoscenze ormai fondamentali. Si può perciò desumere che gli studenti arrivino all'università con competenze

informatiche e digitali insufficienti anche per affrontare compiti basilari come convertire un file o controllare una fonte.

Questo studio vuole quindi indagare le conoscenze digitali di studenti nati dopo il 1993 che attualmente frequentano lauree triennali e magistrali riguardanti materie umanistiche, corsi in cui l'informatica passa spesso in secondo piano sia per quanto riguarda l'interesse degli studenti sia nell'offerta formativa. Inoltre, benché esistano studi sulle competenze digitali degli studenti italiani in generale,<sup>1</sup> non sono a oggi disponibili studi specifici su questa popolazione. Lo studio si propone come indagine conoscitiva per capire come gli studenti percepiscono il proprio livello di competenza digitale e se esiste una discrepanza tra il livello di competenze digitali che gli studenti si auto-assegnano e la conoscenza effettiva.

L'indagine contiene quattro sezioni. La prima comprende alcune domande conoscitive per delineare il profilo dei partecipanti riguardanti, tra le altre:

- anno di nascita
- sesso
- corso di laurea/anno di frequenza
- ateneo di appartenenza
- regione di provenienza
- anni d'uso del computer e della navigazione internet
- dispositivo più usato (computer, tablet, smartphone)
- motivo d'uso più frequente dei dispositivi elettronici (studio, lavoro, svago, comunicazione)
- certificazione ECDL
- completamento di un corso di informatica all'università
- percezione della competenza in confronto ai docenti del liceo e dell'università
- interesse nei confronti dell'informatica

La seconda sezione chiede ai partecipanti di definire la propria competenza in modo molto generico (“Come ritieni la tua competenza informatica?”), collocandola su una scala che va da insufficiente a eccellente.

La terza sezione chiede agli studenti di autovalutare il proprio livello di competenza digitale seguendo il Quadro Europeo delle Competenze Digitali, attualmente utilizzato anche nei modelli di curriculum europeo Europass. In questa domanda sono state inserite le descrizioni complete del Quadro Europeo e gli studenti hanno scelto a quale categoria ritengono di appartenere (utente base, autonomo e avanzato).

La quarta sezione indaga invece alcune competenze degli studenti con quindici domande a risposta multipla che riguardano aspetti pratici e teorici del lavoro con strumenti digitali utili in ambito universitario e professionale, per esempio la gestione e la conversione di file (.rar, PDF, invio dei file...), alcuni aspetti della navigazione su Internet (motori di ricerca, sicurezza, ma anche fonti affidabili per le notizie) e così via. Ogni domanda è stata ideata seguendo le competenze descritte nel Quadro Europeo delle Competenze Digitali ed è strutturata sul modello dei seguenti esempi:

«Per creare un testo in PDF da un documento .doc:

- Lo stampo e lo passo allo scanner
- Uso un programma OCR
- Lo salvo da Word (o altro programma di elaborazione testi)
- Lo apro con Adobe Reader

---

<sup>1</sup> Si vedano, ad esempio, Gui & Argentin 2011, Calvani *et al.* 2012 e Buffardi & Taddeo 2017, che prendono in analisi studenti delle scuole secondarie; l'indagine OCSE-PISA 2015

- Non so»

Oppure:

«Quali di questi sono linguaggi di *mark-up*?

- WordPress
- javascript
- HTML
- XML
- LaTeX
- Excel
- JPEG
- Linux
- Non so»

Le risposte della quarta sezione sono state pesate con un sistema di punteggi e il risultato di ogni rispondente è quindi stato confrontato con le risposte date nella seconda e nella terza sezione del sondaggio per capire se ci sia una discrepanza tra l'autovalutazione delle competenze e la capacità effettiva di utilizzare strumenti digitali.

Il sondaggio è stato completato da più di 300 studenti provenienti da 27 Atenei italiani e raggiunti tramite la condivisione su 46 gruppi universitari di Facebook<sup>2</sup> che riuniscono studenti di atenei, dipartimenti di studi umanistici, ex facoltà di lettere e filosofia, corsi di laurea in lingue e letterature straniere, mediazione linguistica eccetera. Il sondaggio<sup>3</sup> è stato distribuito online e i partecipanti hanno risposto anonimamente<sup>4</sup> alle domande.

I risultati di questo studio saranno utili per tracciare il profilo delle capacità digitali degli studenti universitari di oggi. Inoltre, le risposte date dagli studenti permetteranno di individuare quali conoscenze digitali, a volte date per assodate da docenti e istituzioni, dovrebbero invece essere potenziate durante il percorso scolastico o universitario per fornire agli studenti strumenti adeguati ad affrontare il loro percorso accademico e, più avanti, quello professionale.

## RINGRAZIAMENTI

Ringrazio Luca Tringali, programmatore, per l'aiuto nell'elaborazione delle domande e dei risultati.

## BIBLIOGRAFIA

- [1] Buffardi A., Taddeo G. 2017. *The Web 2.0 Skills of Italian Students: An Empirical Study in Southern Italy*. Italian Journal of Sociology of Education, 9(1), pp. 45-76.
- [2] Calvani A., Fini A., Ranieri M., Picci P. 2012. Are young generations in secondary school digitally competent? A study on Italian teenagers. *Computers & Education*, 58(2), pp. 797-807.

---

<sup>2</sup> L'elenco completo dei gruppi Facebook può essere consultato su <https://pastebin.com/ggBkg5h7> (consultato il 3 gennaio 2019).

<sup>3</sup> Il sondaggio è disponibile su <http://www.zorbaproject.org/informaggio/> (consultato il 3 gennaio 2019).

<sup>4</sup> Per garantire l'anonimato dei partecipanti, le risposte sono state registrate criptando l'IP dei partecipanti con algoritmo di hash MD5.

- [3] Gui M., Argentin G. 2011. Digital skills of internet natives: Different forms of digital literacy in a random sample of northern Italian high school students. *New media & society*, 13(6), pp. 963-980.
- [4] OCSE PISA (2015) *Students, Computers and Learning: Making the Connection. Country note: Italy* <http://www.oecd.org/pisa/keyfindings/PISA-2012-students-computers-italy.pdf> (consultato il 18 ottobre 2018)
- [5] Prensky, M. 2001. Digital natives, digital immigrants part 1. *On the horizon*, 9(5), pp. 1-6.
- [6] Quadro Europeo delle Competenze Digitali <https://europass.cedefop.europa.eu/it/resources/digital-competences> (consultato il 26 ottobre 2018)

# Crowdsourcing Peer Review in the Digital Humanities?

Michael Soprano, Stefano Mizzaro

Department of Mathematics, Computer, and Physical Science, University of Udine, Udine, Italy  
michael.soprano@outlook.com, mizzaro@uniud.it

## ABSTRACT

We propose an alternative approach to the standard peer review activity that aims to exploit the otherwise lost opinions of readers of publications which is called Readersourcing, originally proposed by Mizzaro [1]. Such an approach can be formalized by means of different models which share the same general principles. These models should be able to define a way, to measure the overall quality of a publication as well the reputation of a reader as an assessor; moreover, from these measures it should be possible to derive the reputation of a scholar as an author. We describe an ecosystem called Readersourcing 2.0 which provides an implementation for two Readersourcing models [2, 3] by outlining its goals and requirements. Readersourcing 2.0 will be used in the future to gather fresh data to analyze and validate.

## KEYWORDS

Crowdsourcing, Peer Review, Scholarly Publishing.

## INTRODUCTION

Peer review is at the basis of the scholarly publishing process, the main mechanism to spread scientific knowledge. In it, a scientific article written by some authors is judged and rated by colleagues of the same degree of competence.

Although peer review is a well-established and useful a priori mechanism to ensure the quality of scientific publications, it is not free from problems, related to the process itself and to the malicious behavior of some stakeholders. In some cases reviewers cannot correctly evaluate a publication, e.g., when the paper reports data from an experiment which is long and complex and, therefore, not replicable by the reviewer itself; thus, an act of faith (that the author is honest) is sometimes required [4]. According to a recent article published in *The Economist* [8] there are some journals where the peer review activity is not performed, in contrast with what stated by their publishers. This leads researchers and scholars to “inflate” the list of their publications with articles that, probably, would not pass the peer review filter. There are several discussions related to this phenomenon. One of the main causes (although it is not the only one) is to be found in the change of the business model that allows publishers to get a profit. In recent years those publishers have gone from monetization through the resale of subscriptions to readers to a payment request of a publication fee to the authors of articles. These articles can subsequently be read without any payment according to the open access model. This model, therefore, promotes the dissemination of knowledge but, at the same time, risks corrupting it. The article goes on by describing different malicious behaviors adopted by publishers of at least questionable periodicals to appear respectable and trustworthy when in the reality it is not like that at all. All this is also caused by the institutions of the scientific world which seem to worry less about where the financed research is published.

Another issue is discussed by Mizzaro [1] who, also taking into account several of the flaws of peer review that are widely analyzed in the literature, conjectures that reviewers of scientific publications can be seen as a scarce resource which is being exhausted. To support such a thesis Mizzaro describes in detail ten different factors that contribute to it.

Our main goal in this paper is to start a discussion on exploiting Crowdsourcing techniques in the peer review activity, with a focus on the field of digital humanities, even if we are aware that this field may be characterized by further difficulties compared to others. This paper is structured as follows: we briefly describe an alternative approach to the standard peer review activity called Readersourcing (Section 2) along with the main features of its models (Section 3). We also introduce our Readersourcing 2.0 software ecosystem (Section 4), which allows to take advantage of this alternative approach. Section 5 concludes the paper.

## THE READERSOURCING APPROACH

As a solution to the scarcity of reviewers, as well as to other issues of peer review, Mizzaro [1] proposes to take advantage of readers' opinions by outsourcing the peer review activity to the scholarly community and calls this approach *Readersourcing*, as a portmanteau for “crowdsourcing” and “readers”. As hypothesized by other researchers, it can be assumed that readers are a resource of which there is no shortage: they are many more than the reviewers, so if their opinions can be gathered, they might allow to rate publications quality. Although this might seem a radical solution, it is important to remark that: (i) the effectiveness of crowdsourcing has been demonstrated in several cases [9]; (ii) the usage of crowdsourcing in scholarly publishing is being proposed and analyzed for even more radical approaches, for example to outsource some steps of writing of scientific publications [7]; and (iii) similar approaches, suggesting variants and changes to peer review including collaborative reviews and/or a more distributed peer review practice, have already been proposed in the past [5, 6].

A crucial aspect of the Readersourcing approach that must necessarily be addressed consists in providing a mechanism for which being a “good” reviewer is gratifying, in order to encourage readers to express “good” ratings. As a related issue, some stakeholder of the scholarly publishing process can have a malicious behavior. This is reported from multiple sources and it can be caused by different factors. In this paper we do not aim to provide an exhaustive description of all those behaviours and their causes, but some examples may be mentioned.

This approach is not free from problems itself (e.g., lobbies, lazy readers) [3]; although these need to be taken into consideration, we do not have the space to discuss them.

## READERSOURCING MODELS

To outsource the peer review activity to readers, a model of some kind is useful. Different proposals [2, 3] are available in the literature, and share the same general principles. Every publication is characterized by one or more numerical rating, each one provided by a reader, and these ratings are the input data for such models. From these data the model should define a way to measure the overall quality of a publication as well as the reputation of a reader as an assessor; moreover, from these measures it should be possible to derive the reputation of a scholar as an author.

It is easy to understand that Readersourcing models need to define some specific components: (i) how to aggregate the individual ratings that the assessed entity (i.e., a publication) receives into a single index of quality; (ii) from that index, how to compute a single index of reputation for each assessor (i.e., reader); (iii) eventually, how to compute a single index of how much an author is “skilled”. Moreover, this aggregation must be carried out by taking into consideration that not all ratings are equal and that each of them has an intrinsic adequacy (i.e., a measure of how much truthful and unbiased it is). In other words, it has to be possible to

distinguish adequate from inadequate ratings and, then, good from bad assessors and, again, skilled from unskilled authors. Readersourcing models face these issues by exploiting *co-determination algorithms*. In these algorithms the quality of the assessed entities (i.e., publications) is used to estimate the corresponding reputation/ability of the assessors (i.e., readers). A key point of this characterization is that it allows to exploit the Readersourcing approach as a pre-publication replacement or as a post-publication addition to standard peer review activity.

## THE READERSOURCING 2.0 ECOSYSTEM

We are currently refining a software ecosystem called *Readersourcing 2.0*. Our ecosystem exploits modern ICT tools and complies with four main requirements: (i) to provide the reader with a way to rate a publication by expressing a numerical rating in a seamless and effortless way; (ii) to allow readers to review publications in a way which is independent from the used device (hardware or software); (iii) to be able to aggregate the ratings received by a publication according to different Readersourcing models; (iv) to be general, extensible, and easily adaptable to more models.

Our Readersourcing 2.0 ecosystem, currently in alpha release, is composed of three main applications. A first application allows readers to effectively rate the publications. A second one gathers and aggregates all the ratings given by readers. Those two applications are supported by a third one which has the task of carrying out some technical processing not described in this paper. Figure 1 shows part of the interface of the application used by readers to effectively rate publications within our ecosystem.

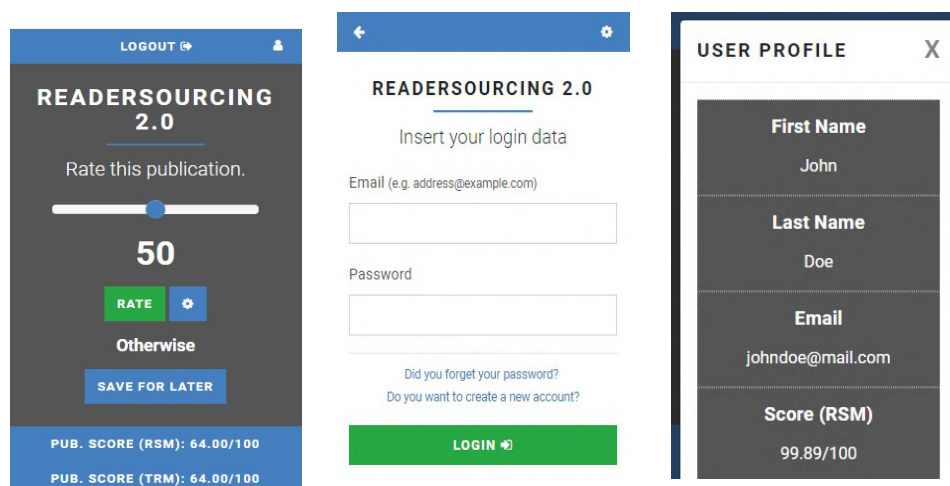


Figure 1: Part of the main user interface used by readers to rate publications within Readersourcing 2.0.

## CONCLUSIONS

Crowdsourcing is a well-established technique which is used in various fields and proposals for its application to peer review already exist. We believe that the Readersourcing approach outlined in Sections 2 and implemented by our Readersourcing 2.0 ecosystem presented in Section 4 can be an interesting alternative, or supplement, to standard peer review activity since it exploits an otherwise lost resource, namely the opinions of readers which can provide a useful contribution. Even if in the field of digital humanities there may be some further specific difficulties due to the more subjective nature of publication quality, it will be

interesting to validate our Readersourcing 2.0 ecosystem on gathered fresh data. The results of these analyses will be the subject of future work.

## REFERENCES

- [1] S. Mizzaro, Readersourcing - A manifesto, *Journal of the American Society for Information Science and Technology*, vol. 63, no. 8, pp. 1666-1672, 2012.
- [9] L. De Alfaro and M. Faella, TrueReview: A Proposal for Post-Publication Peer Review (White Paper), *Computing Research Repository*, 2016.
- [10] S. Mizzaro, Quality Control in Scholarly Publishing: A New Proposal, *Journal of the American Society for Information Science and Technology*, vol. 54, no. 11, pp. 989-1005, 2003.
- [11] W. Y. Arms, What are the alternatives to peer review? Quality Control in Scholarly Publishing on the Web, *The Journal of Electronic Publishing*, vol. 8, no. 1, 2002.
- [12] J. Akst, I Hate Your Paper. Many say the peer review system is broken. Here's how some journals are trying to fix it, *The Scientist*, vol. 24, no. 8, p. 36+, 2010.
- [13] B. Meyers, Fixing the Process of Computer Science Refereeing, 2010. [Online]. Available: <https://cacm.acm.org/blogs/blog-cacm/100030-fixing-the-process-of-computer-science-refereeing/fulltext>.
- [14] Y. Sun, P. Cheng, S. Wang, H. Lyu, M. Lease, I. Marshall and B. C. Wallace, "Crowdsourcing Information Extraction for Biomedical Systematic Reviews," pp. 1-3, 2016.
- [15] The Economist, Some science journals that claim to peer review papers do not do so, *The Economist*, vol. 2018, no. 23.06, 6 2018.
- [16] J. Howe, The Rise of Crowdsourcing, *Wired Magazine*, vol. 14, no. 6, pp. 1-4, 2006.



# Le rappresentazioni digitali e le edizioni di documenti medievali online per la ricerca e la didattica della Diplomatica e della Storia della documentazione. Un caso concreto

Antonella Ambrosio<sup>1</sup>, Maria Rosaria Falcone<sup>2</sup>, Vera Isabell Schwarz-Ricci<sup>3</sup>, Georg Vogeler<sup>4</sup>

<sup>1</sup>Università degli Studi di Napoli Federico II, antonella.ambrosio@unina.it

<sup>2</sup>Università degli Studi di Napoli Federico II, info@monasterium-italia.net

<sup>3</sup>Università degli Studi di Napoli Federico II, vera.schwarz@monasterium-italia.net

<sup>4</sup>Karl-Franzens-Universität Graz, georg.vogeler@uni-graz.at

## ABSTRACT

Nell'ambito di due vasti progetti europei, ENArC (European Network on Archival Cooperation) e co:op (Community as Opportunity. The Creative Archives' and Users' Network), sono state realizzate diverse rappresentazioni digitali di fondi archivistici conservati in Italia e l'edizione digitale dei documenti della prima metà del XIII secolo di S. Maria della Grotta. Entrambi i progetti sono stati compiuti con gli strumenti del portale Monasterium.net. Nel nostro intervento, sulla scorta delle esperienze e delle ricerche svolte, si vuole evidenziare come le rappresentazioni digitali dei documenti medievali e le edizioni digitali possono rappresentare interessanti possibilità per la ricerca scientifica e la didattica della Diplomatica, ed in generale, della Storia della documentazione.

## PAROLE CHIAVE

Documenti medievali online, edizioni critiche digitali, Diplomatica, Storia della documentazione, ambienti collaborativi di descrizione e di ricerca online, codifica XML, Ambienti di apprendimento e tecnologie digitali, équipes di ricerca internazionali e interdisciplinari, progetti europei di Digital Humanities

Nell'ambito di due vasti progetti europei, ENArC (European Network on Archival Cooperation) - EU, Culture Programme, 2007-2013 e co:op (Community as Opportunity. The Creative Archives' and Users' Network) - EU, Creative Europe 2014-2020, sono state realizzate diverse rappresentazioni digitali di fondi archivistici conservati in Italia e l'edizione digitale dei documenti della prima metà del XIII secolo di S. Maria della Grotta, un'abbazia sita a Vitulano, presso Benevento. Entrambi i progetti sono stati compiuti con gli strumenti del portale Monasterium.net, la più ampia risorsa dei documenti pergamenei dell'Occidente europeo oggi esistente sul Web.

Nel nostro intervento, sulla scorta delle esperienze e delle ricerche svolte, si vuole evidenziare come le rappresentazioni digitali dei documenti medievali e le edizioni digitali possono rappresentare nuove e concrete possibilità per la ricerca scientifica e la didattica della Diplomatica, ed in generale, della Storia della documentazione.

La creazione dei nuclei documentari italiani online su Monasterium.net è stata intrapresa nel 2008, in tempi diversi, con svariati obiettivi e differenti modalità da parte di una équipe di studiosi e di archivisti, eterogenea per competenze, provenienza geografica o istituzionale, formatasi grazie a diverse collaborazioni in Italia e all'estero. Essa ha dato luogo ad una risorsa digitale che consta oggi, nel suo complesso, 92580 unità documentarie. Tale tipo di risorse costituiscono, come ha sottolineato recentemente Manfred Thaller, la più lungimirante ed economica forma di valorizzazione e di 'descrizione' delle fonti documentarie, diventando una imprescindibile base per ulteriori e sempre più approfondite chiavi di accesso alla

documentazione (indicizzazioni, descrizioni, registi, edizioni) da effettuare negli anni a venire e con le risorse man mano disponibili.

In questo contesto va inserito anche il lavoro di edizione digitale delle pergamene di S. Maria della Grotta (SMG). L'edizione collaborativa dei documenti di SMG è stata realizzata da un gruppo di studiosi di cui fanno parte Antonella Ambrosio, Giovanni Araldi, Maria Rosaria Falcone, Paola Massa, Vera Schwarz-Ricci, Maria Elisabetta Vendemia, Georg Vogeler. Per la codifica dei 109 documenti privati del periodo 1200-1250 essi hanno utilizzato l'editor online fornito da Monasterium.net, EditMOM3. Monasterium.net è realizzato come un database XML nativo (eXist-db) e utilizza lo standard XML/CEI per la codifica dei documenti medievali. L'ambiente di editing per l'utente è accogliente e riduce le eventuali difficoltà dell'approccio iniziale alla redazione dei contenuti. EditMOM3 è, infatti, un editor WYSIWYM (What You See Is What You Mean) che combina la visualizzazione delle immagini con una maschera strutturata di input e segnala il markup effettuato sul testo inserito tramite virgolette. Per gli utenti più esperti è possibile visualizzare la codifica XML al posto dell'immagine ed effettuare il download del documento in formato xml. Inoltre, la piattaforma permette la condivisione dei propri documenti e l'annotazione delle immagini.

In una prima fase gli editori hanno effettuato la codifica secondo un modello messo a punto per la documentazione privata dell'Italia meridionale (Ambrosio 2018), con l'aiuto di un manuale d'uso sviluppato appositamente e perfezionato con il feedback del gruppo degli editori. Le edizioni digitali e il manuale sono visibili all'indirizzo <<http://monasterium.net/mom/SMG1200-1250/collection>>. Sulla base dei documenti codificati e di un layout di stampa di un'edizione cartacea, in una seconda fase è stata realizzata da Hans Clausen (Center for Information Modelling dell'Università di Graz), con la collaborazione di Vera Schwarz-Ricci, una trasformazione dei files CEI/XML mediante fogli di stile XSLT (eXtensible Stylesheet Language Transformations) in un file XSL-FO (Extensible Stylesheet Language Formatting Objects). Questo file è stato poi convertito con un processore FO in un PDF. Una terza fase è stata dedicata alla stesura, da parte dei curatori, delle ulteriori parti di un'edizione (introduzione, criteri d'edizione, bibliografia e indici) che ha prodotto un PDF complessivo contenente la totalità del lavoro svolto. Questo PDF è stato pubblicato come volume a stampa con codice ISBN (Ambrosio, Schwarz-Ricci, Vogeler 2018a), e sarà distribuito a breve in open access (FEDOA). Quest'ultima soluzione è stata adottata per affrontare il problema del pieno riconoscimento dei prodotti digitali di questo tipo nell'ambito della comunità scientifica di riferimento. Si sottolinea che l'esperimento ha prodotto un workflow replicabile da altri gruppi di ricerca, sia perché i tool utilizzati sono liberamente disponibili sul Web sia perché la documentazione realizzata è disponibile online per i potenziali editori.

La realizzazione delle rappresentazioni digitali delle quali si è trattato sopra e la produzione dell'edizione digitale hanno avuto interessanti ricadute sulla didattica, nella quale fin dal 2008 sono stati utilizzati gli strumenti delle Digital Humanities, anche in collaborazione con le cattedre di Pedagogia dell'Ateneo. Questi ultimi hanno facilitato la produzione di ambienti di apprendimento ideali secondo l'approccio costruttivista che hanno combinato l'insegnamento tradizionale soprattutto con il Learning by doing. Corsi di questo genere sono stati organizzati nell'ambito del corso di Paleografia e di Diplomatica del corso di Laurea in Scienze storiche dell'Università degli Studi di Napoli Federico II e in un Master *post lauream* in Biblioteconomia e Archivistica, attivo nello stesso ateneo dal 2009 al 2011. Si è generato, in questi casi, un circolo virtuoso, da una parte le esperienze di ricerca hanno contribuito a rendere l'ambiente di apprendimento più efficace, dall'altra gli studenti hanno partecipato attivamente ad esso con interventi improntati allo spirito critico durante i corsi, con la partecipazione ai workshop e ai meeting internazionali organizzati nell'ambito della progettualità europea, con la produzione di descrizioni, registi, edizioni, indicizzazioni che

non di rado sono stati pubblicati online. Ciò ha prodotto in essi un senso di appartenenza alla community internazionale che anima il portale Monasterium.net, composta di ricercatori, di studiosi, di archivisti provenienti da diversi paesi europei, potenziando enormemente il loro percorso formativo e la capacità di orientare il loro futuro professionale.

## BIBLIOGRAFIA

- [1] Aigner, T. 2010. «Monasterium.Net - Documenti Europei online». *Archivi. Associazione Nazionale Archivistica Italiana* 5 (2): 123–28.
- [17] Ambrosio, A. 2010. «Il progetto Monasterium in Italia: le prime esperienze a Napoli». *Archivi. Associazione Nazionale Archivistica Italiana* 5 (2): 129–46.
- [18] Ambrosio, A. 2011. «Towards the creation of a learning environment within the monasterium project: teaching experiences of diplomatics». In Aigner, T., S. Hohenbruck, T. Just, e J. Kemper, a c. di. *Archive im Web - Erfahrungen, Herausforderungen, Visionen / Archives on the Web. Experiences, Challenges, Visions*. St. Pölten: Diözesanarchiv St. Pölten, 203–15.
- [19] Ambrosio, A. 2012. «Insegnare la diplomatica con le nuove tecnologie. Potenzialità e spunti di riflessione». In Cherubini, P., e G. Nicolaj, a c. di. *Sit liber gratus, quem servulus est operatus: studi in onore di Alessandro Pratesi per il suo 90° compleanno*. 2 vol. Littera antiqua 19. Città del Vaticano: Scuola vaticana di paleografia, diplomatica e archivistica, 2°:1315–26.
- [20] Ambrosio, A. 2014. «Transnational Virtual Learning Communities and Historical Documents». In M. J. Salamanca López, C. H. Artaza, *Innovación didáctica al servicio del docente y profesional en Ciencias Documentales*, Lima: REDEM, 29–39.
- [21] Ambrosio, A. 2018. «L'edizione critica digitale dei documenti medievali. Le forme degli atti di Octavianus notarius». In Figliuolo, B., R. Di Meglio, e A. Ambrosio, a c. di. *Ingenita curiositas. Studi sull'Italia medievale per Giovanni Vitolo*, 3:1153–73. Battipaglia (Sa): Laveglia & Carlone.
- [22] Ambrosio, A. «Digital critical editions of medieval documents on Monasterium.Net, in *Online edition of medieval archives: issues, methodology and challenges*». In Atelier de recherche sur les textes médiévaux - ARTEM publiée par l'Atelier Diplomatique (CRULH), Brepols editions (in corso di stampa).
- [23] Ambrosio, A., G. Aiello, e M. R. Falcone. 2012. «The Icarus Didactics Group. A challenge of teaching method of Diplomatics, Palaeography and Medieval History». In Mokryš, M., e A. Lieskovský, a c. di. *ARSA 2012. Proceedings in Advanced Research in Scientific Areas: The 1st Virtual International Conference*. Žilina: EDIS - Publishing Institution of the University of Žilina, 701–5.
- [24] Ambrosio, A. M. Striano, C. Freda, S. Fiorentino, e L. Aiello. 2012. «Teaching Diplomatics in 2.0 Web Environments: An Innovative Experience to Promote Interaction Among Students From Different Countries and With Different Learning Needs». In Beldhuis, H., a c. di. *Proceedings of the 11th European Conference on e-Learning: ECEL*, Reading: Academic Publishing International, 15-9.
- [25] Ambrosio, A., V. I. Schwarz-Ricci, e G. Vogeler, a c. di. 2018a. *I documenti dell'abbazia di S. Maria della Grotta di Vitulano (BN). 1200-1250. Edizione a stampa* (con le edizioni di A. Ambrosio, G. Araldi, M. R. Falcone, P. Massa, V. I. Schwarz-Ricci, M. E. Vendemia, G. Vogeler). Battipaglia (Salerno): Laveglia & Carlone.
- [26] Ambrosio, A., V. I. Schwarz-Ricci, e G. Vogeler, a c. di. 2018b. *I documenti dell'abbazia di S. Maria della Grotta di Vitulano (BN). 1200-1250* (con le edizioni di A. Ambrosio, G. Araldi, M. R. Falcone, P. Massa, V. I. Schwarz-Ricci, M. E. Vendemia, G. Vogeler), versione digitale, <<http://monasterium.net/mom/SMG1200-1250/collection>>

- [27] Ansani, M. 2006. «Edizione digitale di fonti diplomatiche: esperienze, modelli testuali, priorità». *Reti medievali* 7, <<http://www.rmoa.unina.it/1896/1/140-380-1-PB.pdf>>.
- [28] Boot, P., A. Cappellotto, W. Dillen, F. Fischer, A. Kelly, A. Mertgens, A.-M. Sichani, e E. Spadini & D. van Hulle, a. c. di. 2017. *Advances in Digital Scholarly Editing. Papers presented at the DiXiT conferences in The Hague, Cologne, and Antwerp*. Leiden: Sidestone, <<https://www.sidestone.com/books/advances-in-digital-scholarly-editing>>.
- [29] Burkard, B. 2009. «EditMOM - ein spezialisiertes Werkzeug zur kollaborativen Urkunden-Erschließung». In Vogeler, G., a. c. di. *Digitale Diplomatie. Neue Technologien in der historischen Arbeit mit Urkunden*. Archiv für Diplomatie. Beiheft 12. Köln, Weimar, Wien: Böhlau, 255–70.
- [30] *Charters Encoding Initiative (CEI)*, <<http://www.cei.uni-muenchen.de/index.php>>.
- [31] co:op (Community as Opportunity. The Creative Archives' and Users' Network) - EU, Creative Europe 2014-2020 (<<http://www.coop-unina.org/>>).
- [32] ENArC (European Network on Archival Cooperation) - EU, Culture Programme, 2007-2013 (<<http://www.recruitdigitaldoc.org/>>).
- [33] Falcone, M. R. 2014. «Il portale Monasterium.net. Documenti in rete e archivi digitali». In Caldelli, E., M. Maniaci, e S. Zamponi, a c. di. *Manuscript Digitization and on Line Accessibility. What's Going on? International Workshop* (Roma, Biblioteca Vallicelliana, 23 ottobre 2014), Digitalia, 2. Roma: ICCU. 67-77, <<http://digitalia.sbn.it/article/view/1532/1032>>.
- [34] Fischer, F. 2013. «All texts are equal, but...: textual plurality and the critical text in digital scholarly editions». *Variants*, n. 10: 77–91, <[https://doi.org/10.1163/9789401209021\\_007](https://doi.org/10.1163/9789401209021_007)>.
- [35] Federico II Open Archive, l'archivio istituzionale dei documenti digitali dell'Università degli Studi di Napoli Federico II (FEDOA), <<http://www.fedoa.unina.it/>>.
- [36] Monasterium.net, <<http://www.monasterium.net>>.
- [37] Pierazzo, E. 2015. *Digital Scholarly Editing: Theories, Models and Methods*. Farnham.
- [38] Robinson, P. 2004. «Where We Are with Electronic Scholarly Editions, and Where We Want to Be». *Jahrbuch für Computerphilologie* 5: 123–43.
- [39] Sahle, P. 2013. *Digitale Editionsformen: zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels*, 3 vol. Schriften des Instituts für Dokumentologie und Editorik 7-9. Norderstedt: BoD.
- [40] Sahle, P. 2016. «What is a scholarly digital edition (SDE)?» In Driscoll, M. J., e E. Pierazzo, a c. di. *Digital Scholarly Editing: Theories and Practices*. Digital Humanities Series 4. Open Book Publisher: Cambridge, 19–39, <<http://dx.doi.org/10.11647/OBP.0095.02>>.
- [41] Siemens, R., M. Timney, C. Leitch, C. Koolen, e A. Garnett. 2012. «Toward Modeling the “social” Edition: An Approach to Understanding the Electronic Scholarly Edition in the Context of New and Emerging Social Media». *Literary and Linguistic Computing* 27 (4): 445–61. <<https://doi.org/10.1093/lc/fqs013>>.
- [42] Thaller, M. 2018. «Archivi e edizioni: alcune promesse mantenute. Un punto di vista personale». In Ambrosio, A., V. I. Schwarz-Ricci, & G. Vogeler, a c. di. *I documenti dell'abbazia di S. Maria della Grotta di Vitulano (BN). 1200-1250. Edizione a stampa* (con le edizioni di A. Ambrosio, G. Araldi, M. R. Falcone, P. Massa, V. I. Schwarz-Ricci, M. E. Vendemia, G. Vogeler). Battipaglia (Salerno): Laveglia & Carlone, XI–XIV.
- [43] Vasold, G. 2014. «Progressive Editionen als multidimensionale Informationsräume». In Ambrosio, A., S. Barret, e G. Vogeler, a c. di. *Digital diplomatics. The computer as a tool for the diplomatist? Archiv für Diplomatie*. Beiheft 14. Köln, Weimar, Wien: Böhlau, 75–88, <<http://dx.doi.org/10.7788/boehlau.9783412217020.75>>.

- [44] Vogeler, G. 2004. «Ein Standard für die Digitalisierung mittelalterlicher Urkunden mit XML. Bericht von einem internationalen Workshop in München 5./6. April 2004». *Archiv für Diplomatik* 50: 23–34, <<https://doi.org/10.7788/afd.2004.50.jg.23>>.
- [45] Vogeler, G. 2004. «Uno standard per la digitalizzazione dei documenti medievali con XML. Cronaca di un Workshop internazionale: Monaco 5-6 aprile 2004». *Scrineum* 2: 241–55, <<http://dx.doi.org/10.13128/Scrineum-12104>>.
- [46] Vogeler, G. 2005. «Towards a Standard of Encoding Medieval Charters with XML». *Literary and Linguistic Computing* 20 (3): 269–80, <<https://doi.org/10.1093/lc/fqi031>>.
- [47] Vogeler, G. 2006. «Charters Encoding Initiative (CEI). Zu Möglichkeiten der Integration mit Hilfe eines Standards für Urkundendigitalisierung». In Aigner, T., e K. Winter, a c. di. *Alte Archive - Neue Technologien. Old Archives - New Technologies*. St. Pölten: Diözesanarchiv St. Pölten, 181–98.

# La monografia digitale: una riflessione sulle implicazioni dell'innovazione testuale

Alessandra Di Tella

Università di Bologna, Italy, [alessandra.ditella2@unibo.it](mailto:alessandra.ditella2@unibo.it)

## ABSTRACT

Prendendo in esame il campo dell'editoria accademica, considerato come luogo di incontro dei due paradigmi della critica culturale umanistica e delle Digital Humanities, il presente contributo intende proporre una riflessione sulle prospettive del nuovo genere della monografia digitale. In particolare, partendo dall'esperienza maturata alla Brown University durante un periodo di studio e collaborazione nell'ambito del progetto Digital Publications Initiative, si vogliono portare all'attenzione potenzialità e problemi legati all'innovazione del testo e alle sue implicazioni nello sviluppo del genere in questione. Basato sui principi della *digital rhetoric* e del *design as argument*, il testo è interamente pensato e progettato per l'ambiente digitale; si caratterizza per una struttura modulare che favorisce la multimodalità e incorpora al suo interno tools provenienti dal paradigma delle Digital Humanities. Distaccandosi definitivamente dalla tradizione a stampa e ampliando notevolmente le possibilità di espressione della complessità della ricerca, l'innovazione testuale comporta implicazioni di vario tipo. Supportata da esempi e materiali acquisiti, l'analisi si focalizza su tre punti principali: la ridefinizione del workflow editoriale e delle fasi di progettazione, composizione e revisione del testo; la comparsa di un forte grado di co-autorialità, dovuto all'eterogeneità delle competenze necessarie allo sviluppo dell'*argument*; l'utilizzo e la scelta degli ambienti e degli strumenti deputati alla scrittura digitale. Ne risulta un panorama ad alto tasso di sperimentazione tecnica e concettuale. Problematizzando l'esperienza svolta, si vuole dimostrare come l'assenza di una prassi consolidata rappresenti, per i progetti in corso, l'opportunità di procedere empiricamente verso la creazione di un metodo: nel tentativo di fornire soluzioni e risposte dettate dalle singole contingenze, si fa strada la possibilità di delineare un modello per la futura diffusione e 'sistematizzazione' del genere. Si accenna, infine, a come la stessa strategia 'a posteriori' si applichi non solo allo sviluppo testuale, ma anche a questioni di processo più tecniche come valutazione, pubblicazione, distribuzione e conservazione della monografia digitale.

## PAROLE CHIAVE

Editoria accademica, monografia digitale, innovazione testuale, autorialità, testo digitale, workflow editoriale.

## LA MONOGRAFIA DIGITALE TRA TECNOLOGIA E STUDI CULTURALI

Le Digital Humanities hanno trasformato il paradigma della ricerca umanistica grazie allo sviluppo di strumenti e metodi a supporto della produzione, della trasmissione e della fruizione del sapere. Lo scambio non è stato unilaterale, poiché la cultura umanistica tradizionale e la critica culturale costituiscono una solida base su cui costruire tools e metodologie digitali.

Il presente contributo prende in esame il campo dell'editoria accademica; territorio ancora poco esplorato da questo punto di vista, ma molto fertile per il dialogo tra tecnologia e studi culturali.

In particolare, il genere della monografia accademica di ambito umanistico, ormai da decenni in crisi (Chodorov 1999, Darnton 1999, Thompson 2005), sta subendo un'importante trasformazione digitale (Riva 2017) che obbliga a riconfigurare *authoring* e *reading* (Balsamo 2018), dando vita a un nuovo modo di costruire e disseminare la conoscenza scientifica. In questo senso l'apporto delle Digital Humanities e della Cultura Digitale (Gere 2009) gioca un ruolo fondamentale nel distacco dal modello a stampa (con cui in generale si intende una tradizione fortemente ancorata alla scrittura lineare, sia essa supportata da un libro cartaceo o da un e-book), condizione imprescindibile per il ripensamento del genere.

Questo contributo trae origine da un percorso di ricerca svolto all'interno di un progetto di dottorato *in fieri*, e si rifà in particolare all'esperienza maturata alla Brown University durante un periodo trimestrale di studio e collaborazione alla Digital Publishing Initiative<sup>1</sup> (condotta nell'ambito del più ampio progetto Monograph Publishing in the Digital Age<sup>2</sup>, finanziato dalla Mellon Foundation). La collaborazione ha riguardato in particolare lo sviluppo della monografia *Italian Shadows*, tuttora in fase di produzione, frutto del lavoro del Prof. Massimo Riva e del suo team. Partendo dall'analisi di materiali, esempi e problemi rilevati, si vuole proporre un percorso di riflessione critica basato sul tipo di testo che caratterizza la monografia digitale e sulle implicazioni che il testo stesso comporta nel processo compositivo (Journet 2012).

## IMPLICAZIONI DELL'INNOVAZIONE TESTUALE

Basato sui principi della *digital rhetoric* e del *design as argument* (Eyman 2015), il nuovo tipo di testualità è *born-digital*: interamente pensato e progettato per l'ambiente digitale. La sua struttura modulare predilige l'alternanza multimodale alla predominanza del testo scritto, e guida l'utente attraverso un'esperienza di lettura non lineare e interattiva. L'eventuale utilizzo di tools provenienti dall'orizzonte delle Digital Humanities (VR reconstruction, GIS mapping, strumenti di text-mining, solo per citarne un paio) risulta parte integrante della composizione testuale e fornisce nuove possibilità di esprimere la complessità del lavoro di ricerca.

Come sperimentato sul campo durante l'esperienza statunitense, l'innovazione in termini di forma del testo interessa la produzione della monografia su vari livelli. Si vuole focalizzare l'analisi, supportata da materiali ed esempi, in particolare su:

a) workflow editoriale – le fasi di progettazione, composizione e revisione del testo, rispetto ad una scrittura di tipo tradizionale, sono meno distinte e più soggette a reciproca sovrapposizione (sia sincronicamente che diacronicamente). La variante tecnologica, infatti, amplia le potenzialità dell'*argument* e pone continuamente gli autori di fronte a scelte, aggiustamenti, ridefinizioni, cambi di rotta. In altre parole, essa costringe a un più alto livello di sperimentazione.

b) autorialità e competenze – la necessità di molteplici specializzazioni, complementari ma egualmente concorrenti allo sviluppo del progetto, mette in discussione la tradizionale concezione di autorialità. Ideato a partire dal lavoro di ricerca dell'autore, l'*argument* prende forma e si struttura grazie al contributo collettivo delle diverse competenze coinvolte: web-developer, designer e editor, ad esempio, sono alcune delle figure professionali che collaborano sinergicamente con l'autore.

c) ambiente e strumenti di scrittura – la scelta, da calibrare a seconda del tipo di contenuto che si vuole sviluppare, può ricadere sulla creazione di un sito web autonomamente implementato oppure sull'utilizzo di una delle (poche) piattaforme esistenti che si prestano alle necessità della scrittura digitale.

---

<sup>1</sup> <https://library.brown.edu/create/digitalpublications/>

<sup>2</sup> <https://mellon.org/resources/shared-experiences-blog/monograph-publishing-digital-age/>

Si confà a questo tipo di operazione la piattaforma Scalar<sup>3</sup> (sviluppata dall'Alliance for Networking Visual Culture), uno degli strumenti più performanti a disposizione dei 'pionieri del genere', il cui punto di forza sta in una certa facilità di utilizzo, che nulla sottrae all'efficacia delle pubblicazioni ma permette anche a un utilizzatore non specialista di sviluppare un lavoro soddisfacente e in linea con i principi della scrittura e della lettura digitali. Di contro, grazie ad una buona flessibilità, Scalar offre ad un utente esperto possibilità di mashup e personalizzazione (costruzione di visualizzazioni e interfacce tramite API, per esempio).

## **COSTRUIRE UN MODELLO TRAMITE L'ESPERIENZA**

Riportando e problematizzando l'esperienza fatta, si vuole dimostrare come non esista, ad oggi, una prassi consolidata di sviluppo del testo della monografia digitale. Trattandosi di un genere ancora in via di sperimentazione, i progetti in corso si trovano a fronteggiare esigenze e istanze inedite. Nel tentativo di fornire soluzioni e risposte dettate dalle singole contingenze, c'è la possibilità di delineare un modello metodologico generale che funga da guida per la futura diffusione e 'sistematizzazione' del genere. In quest'ottica si approfondiranno, in sede di presentazione, alcune delle metodologie/soluzioni applicate in *Italian Shadows* per far fronte ai punti di cui sopra, in particolare:

- a) la convergenza delle fasi di scrittura, implementazione, revisione del testo e ideazione della *narrative*.
- b) la strategia di collaborazione tra le tre principali *expertise* del team: di ricerca, editoriale, tecnica.
- c) l'approccio ai limiti e ai punti di forza della piattaforma Scalar.

La medesima strategia 'a posteriori' è applicata per far fronte alle questioni editoriali di pubblicazione, valutazione, distribuzione, conservazione, ecc. Sebbene esse rappresentino l'ampiezza e la complessità dell'argomento, in questa sede vi si può solo accennare, rimanendo consapevoli della necessità di estendere la discussione e il dialogo ai vari attori del processo: editori, ricercatori, atenei, centri di ricerca.

## **BIBLIOGRAFIA ESSENZIALE**

- [1] Ball, C. E. e Eyman, D. 2015. *Editorial Workflows for Multimedia-Rich Scholarship*. In: *The Journal of Electronic Publishing* 18 (4). <http://dx.doi.org/10.3998/3336451.0018.406>.
- [2] Balsamo, A. 2018. "Making Meaning, Making Culture: How to Think about Technology and Cultural Reproduction". In: *The Routledge Companion to Media Studies and Digital Humanities*. A cura di Sayers J. Londra: Routledge.
- [3] Chodorow, S. 1999. *Specialized Scholarly Monograph in Crisis or How Can I Get Tenure if you Won't Publish My Book?*. Washington DC: Association of Research Libraries.
- [4] Darnton, R. 1999. *The New Age of the Book*. In: *New York Review of Books* 46 (5).
- [5] Eyman, D. e Ball, Cheryl E. 2015. "Digital Humanities Scholarship and Electronic Publication". In: *Rhetoric and the Digital Humanities*. A cura di Ridolfo J. Chicago: University of Chicago Press. Pp. 65-79.
- [6] Finkelstein, D. e McCleery A. (a cura di). 2002. *The book history reader*. Londra: Routledge.

---

<sup>3</sup> <http://scalar.me/>



- [7] Journet, D., Ball, C.E. e Trauman, R. (a cura di). 2012. *The New Work of Composing*. Logan: Computers and Composition Digital Press/Utah State University Press.  
Disponibile qui: <http://ccdigitalpress.org/nwc>.
- [8] Marx, I. e Smith, M. R. (a cura di). 1994. *Does Technology Drive History? The Dilemma of Technological Determinism*. Cambridge: The MIT Press.
- [9] Maxwell, J. W., Bordini, A. e Shamash, K. 2017. *Reassembling Scholarly Communications: An Evaluation of the Andrew W. Mellon Foundation's Monograph Initiative* (Final Report, May 2016). In: *The Journal of Electronic Publishing* 20 (1).  
<http://dx.doi.org/10.3998/3336451.0020.101>.
- [10] Riva, M. 2012. *Liquid/Cloudy/Foggy: For a Critique of Fluid Textuality*. In: *Humanist Studies & the Digital Age* 2 (1). <http://dx.doi.org/10.5399/uo/hsda.2.1.3016>.
- [11] Riva, M. 2017. An Emerging Scholarly Form: the Digital Monograph. In: *Digitcult* 2 (3). Pp. 63-74.  
<http://dx.doi.org/10.4399/97888255099087>.
- [12] Gere, C. 2009. *Digital Culture*. Londra: Reaktion books.
- [13] Pentz, E., Borthwick, I. e Campbell, R. 2012. *Academic and Professional Publishing*. Amsterdam: Elsevier.
- [14] Thompson, J. 2005. *Books in the Digital Age: The Transformation of Academic and Higher Education Publishing in Britain and the United States*. Cambridge: Polity.
- [15] Williams, P., Stevenson, I., Nicholas, D., Watkinson A. e Rowlands, I. 2009. *The Role and the Future of the Monograph in Arts and Humanities*. In: *Aslib Proceedings* 61 (1).  
<https://doi.org/10.1108/00012530910932294>.

# Manzoni digitale: un laboratorio tra didattica e ricerca

Beatrice Nava

Università di Bologna, Italy, [beatrice.nava2@unibo.it](mailto:beatrice.nava2@unibo.it)

## ABSTRACT

Il presente contributo si propone di descrivere un caso concreto di didattica nell'ambito delle DH, attualmente in corso di svolgimento presso il dipartimento FICLIT dell'Università di Bologna: una serie di laboratori per la marcatura XML/TEI di testi di Alessandro Manzoni, inserita all'interno del PRIN *Manzonionline: carte, libri, edizioni, strumenti*. Il progetto è dedicato alla creazione di un ambiente digitale nel quale saranno resi disponibili e consultabili tutti i materiali relativi al lavoro dell'autore, dai manoscritti alle biblioteche, dalla bibliografia critica alle opere. L'obiettivo è dunque quello di presentare un esperimento di didattica integrato all'interno di un progetto di ricerca e di analizzare, attraverso la somministrazione di un breve questionario di valutazione, la risposta degli studenti di due distinti laboratori, svolti seguendo metodologie didattiche differenti. In tal modo si potrà testare quale approccio didattico risulti non solo più vicino alle esigenze degli studenti, ma anche più funzionale all'erogazione delle competenze di base previste dall'insegnamento e a una reale comprensione del funzionamento, delle potenzialità e delle criticità del vocabolario TEI per la codifica di testi letterari moderni. Si analizzeranno anche limiti e benefici dell'approccio collaborativo tra didattica e ricerca adottato per fornire competenze teoriche e pratiche di base per l'uso del linguaggio XML/TEI e collocare gli studenti in un progetto reale e percepito come utile per la comunità scientifica, contesto ideale per favorire una maggiore consapevolezza dell'effettivo funzionamento di un progetto di ricerca.

## PAROLE CHIAVE

Codifica di testi moderni, Alessandro Manzoni, Laboratorio didattico, Integrazione didattica-ricerca, Marcatura XML/TEI

## INTRODUZIONE

A fronte di una sempre più universalmente riconosciuta importanza e autonomia disciplinare delle DH anche nel panorama didattico nazionale, ancora relativamente pochi<sup>4</sup> rimangono i casi di master universitari o corsi di laurea interamente dedicati a questo settore: fanno scuola i due casi di Bologna e Pisa, con i corsi, rispettivamente, in *Digital Humanities and Digital Knowledge* (corso di laurea magistrale, LM-43) e in Informatica Umanistica (corso sia triennale che magistrale: L-10 e LM-43). Tuttavia, è pur vero che in varie Università iniziano a farsi strada numerosi seminari/laboratori dedicati all'approfondimento di specifici aspetti ascrivibili a questa vasta ed eterogenea famiglia di discipline.

All'interno di un panorama ancora caratterizzato, dunque, da un certo grado di sperimentazione e spontaneismo, misurabile sia nella difformità dei contenuti erogati, che in una collocazione e riconoscimento istituzionale non ancora del tutto definiti, si corre spesso il rischio di avviare singoli esperimenti didattici privi di una vera e propria contestualizzazione che li leghi al corso di studi che li ospita o promuove. Ne consegue uno sperimentalismo

---

<sup>4</sup> Si veda la sintetica lista fornita dal sito Filologia Digitale Verona, del Dipartimento di Lingue e Letterature straniere dell'Università di Verona: <http://filologiadigitale-verona.it/uncategorized/corsi-dh-nelle-universita-italiane/>

spesso poco organizzato e coordinato con l'offerta didattica generale, che ha l'effetto di disorientare o allontanare gli studenti, i quali non riescono a cogliere il valore del singolo corso di ambito DH entro il proprio percorso di studi.

Scopo quindi dei laboratori per la marcatura XML/TEI di testi manzoniani avviati presso il dipartimento FICLIT dell'Università di Bologna è fornire, all'interno del corso di Laurea Magistrale in Italianistica, uno spazio in cui sperimentare concretamente la diretta continuità tra gli studi umanistici tradizionali e quelli di dominio delle discipline digitali, in modo da mostrare agli studenti come l'acquisizione di nuove competenze in questo settore possa arricchire l'esperienza di un tradizionale studioso di letteratura e filologia, fornendo nuovi strumenti di analisi, interpretazione e fruizione dei testi letterari. A ulteriore sostegno di questa volontà, i laboratori – coordinati dalla professoressa Paola Italia – sono stati inquadrati nell'ambito di un progetto PRIN (PI Giulia Raboni, Unipr) dedicato all'opera di Alessandro Manzoni, all'interno del quale l'unità bolognese (ru Paola Italia, Unibo) ha l'obiettivo di realizzare un archivio delle più aggiornate e affidabili edizioni manzoniane, lavoro che comprende la digitalizzazione, mediante codifica XML/TEI, di edizioni scelte e controllate dei testi dell'autore.

Gli studenti che partecipano a questo ciclo di laboratori sono dunque inseriti entro un vero progetto di ricerca e il loro contributo avrà una ricaduta concreta sull'ambiente digitale in via di costruzione; ciò, oltre ad essere un aspetto intrinsecamente motivante, favorisce una maggiore comprensione delle dinamiche che caratterizzano un progetto di ricerca, della necessità di apporti multidisciplinari e a diverso livello di competenza e, soprattutto, dell'utilità per la comunità scientifica del lavoro svolto durante i laboratori e, più in generale, delle conoscenze e competenze acquisite, anche nell'ottica dell'inserimento in ambiti professionali e occupazionali sempre più attrattivi e fiorenti.

## **METODOLOGIE DIDATTICHE**

Finora sono stati attivati due laboratori, tenuti tra il 20 marzo e il 15 maggio e tra il 15 ottobre e il 17 dicembre 2018. Entrambi sono stati pensati per gli studenti del corso di Laurea Magistrale in Italianistica, ma sono stati ammessi anche studenti di provenienza diversa<sup>5</sup>, a riprova dell'attrattività trasversale di questo tipo di competenze.

Scopo dell'insegnamento è quello di fornire una panoramica sulla situazione attuale dell'editoria digitale, con particolare attenzione allo stato dell'arte in materia di *digital scholarly editing* e alle criticità relative all'edizione digitale dei testi letterari moderni, con l'obiettivo concreto di portare lo studente a un utilizzo autonomo e consapevole del linguaggio XML/TEI. Per questa ragione si è assegnata a ciascun iscritto la marcatura di un'opera manzoniana (o di parte di essa, per i casi di testi molto corposi), da inserire – una volta corretta – nella piattaforma *Manzonionline*.

In entrambi i casi il corso è stato introdotto da due lezioni, tenute dalla professoressa Paola Italia, dedicate alla presentazione del progetto PRIN e alle principali questioni generali relative alla filologia d'autore, ivi compresi i problemi della collazione (e della scelta delle edizioni da fornire sul portale). A questa prima introduzione hanno fatto seguito tre incontri, svolti anch'essi in forma di tradizionale lezione frontale, con l'uso di presentazioni poi fornite agli studenti, relativi alle principali questioni teoriche e metodologiche da affrontare nel corso (introduzione all'editoria digitale; concetti di dato computabile e di codifica; dato come informazione; tipologie di edizione digitale; potenzialità e criticità delle edizioni digitali; linguaggi di markup; concetto di metalinguaggio; SGML, XML, DTD e XML Schema; introduzione alla Text Encoding Initiative; TeiHeader e body; i moduli TEI). Terminata

---

<sup>5</sup> Al primo laboratorio hanno partecipato anche due studentesse di Lettere Classiche, mentre il secondo laboratorio è stato frequentato anche da una dottoranda e da due assegniste di ricerca.

questa prima parte, sono state avviate le esercitazioni laboratoriali, con la marcatura, condivisa prima e autonoma poi, di semplici testi d'esempio, per poi passare alla codifica del testo manzoniano assegnato. La verifica finale delle competenze è stata svolta mediante colloquio orale vertente sulla parte teorica, sulla discussione degli eventuali errori riscontrati nel testo marcato e sulla possibile marcatura alternativa di specifici fenomeni testuali.

Tenuto fisso questo impianto metodologico di base, i due laboratori sono stati poi impostati diversamente per quanto riguarda l'organizzazione del calendario, gli strumenti forniti e la modalità pratica di approccio alla marcatura del testo. Infatti, per il primo laboratorio sono stati organizzati due incontri settimanali della durata di due ore ciascuno, per un totale di 30 ore di lezione in aula, mentre nel secondo si è scelto di proporre lezioni più lunghe, di tre ore ciascuna, ritenendo questa durata più confacente al tipo di lavoro da svolgere. Gli strumenti didattici presentati, in particolare le Guidelines TEI P5 e TEI by Example, sono stati i medesimi, ma durante il primo corso sono stati proposti come supporto didattico facoltativo, al quale lo studente poteva scegliere autonomamente di riferirsi, mentre nel secondo caso si è favorito un effettivo utilizzo in aula. Un'altra importante differenza ha riguardato la definizione del modello teorico di marcatura da applicare al testo manzoniano: nel primo caso si è fornito un modello predefinito, concordato con una discussione corale in classe, nel secondo, invece, è stato chiesto a ciascuno studente di individuare i fenomeni più interessanti da codificare nel testo assegnato e le possibili soluzioni di codifica, attraverso l'uso autonomo delle Guidelines TEI. In questo modo l'adeguamento delle varie proposte al modello scelto dal gruppo di ricerca PRIN è stato pensato come un momento successivo e slegato dall'interpretazione del testo e dalla scelta della codifica da applicare, con un guadagno in termini di autonomia dello studente, a scapito di un maggiore sforzo previsto in fase di uniformazione e correzione dei file prodotti.

## I QUESTIONARI DI VALUTAZIONE

Tutti i correttivi sopra accennati sono stati applicati con l'intenzione di porre maggiormente l'accento sull'autonomia dello studente, invitato a cercare da solo possibili soluzioni di codifica, così da poter sperimentare direttamente la vastità delle possibilità offerte dal vocabolario TEI. Ciò anche con l'obiettivo di fornire conoscenze e, soprattutto, strumenti per un futuro approfondimento della materia e per un più agevole adeguamento di quanto appreso a eventuali interessi diversi dallo specifico caso manzoniano proposto.

Alla luce di queste modifiche nell'approccio didattico e per verificarne l'effettiva ricaduta, è stato preparato un semplice questionario da somministrare agli studenti di entrambi i laboratori: i quesiti mirano a indagare direttamente l'opinione riguardo ai singoli aspetti in cui si è articolata la didattica, in particolare relativamente ai punti maggiormente divergenti nei due corsi.

Le domande proposte, alle quali lo studente può rispondere decisamente no, più no che sì, più sì che no e decisamente sì, sono le seguenti:

### **Organizzazione calendario:**

- 1) Hai trovato la durata della singola lezione adeguata al lavoro richiesto?
- 2) Hai trovato eccessivo il lavoro domestico necessario a completare la marcatura dei testi affidati?
- 3) Hai portato a termine la marcatura di tutti i testi assegnati?

### **Strumenti forniti:**

- 4) Hai trovato utili a fini didattici gli strumenti presentati a lezione?

- 5) Hai utilizzato concretamente gli strumenti forniti a lezione (Guidelines, Tei by Example..)?  
6) Hai trovato utili per un futuro utilizzo autonomo gli strumenti proposti?

**Definizione del modello di marcatura:**

- 7) Ritieni di aver maturato competenze utili per un futuro inserimento lavorativo/prosecuzione degli studi?  
8) Ritieni di aver acquisito la capacità di affrontare autonomamente la codifica di un testo in futuro?  
9) Definiresti la codifica un processo meccanico o interpretativo, in base alla sola esperienza concreta maturata durante il laboratorio?

**RISULTATI ATTESI E CONCLUSIONI**

Considerando lo spostamento dell'approccio didattico del secondo laboratorio verso un coinvolgimento più attivo dello studente, ci si aspettano risposte più positive dagli iscritti al secondo laboratorio, in particolare alle domande relative all'uso degli strumenti e alla definizione del modello di marcatura. D'altro canto, considerando che il primo corso è stato maggiormente focalizzato sul risultato concreto da ottenere, ci si aspetta che gli studenti di questo laboratorio portino a compimento più velocemente la marcatura del testo/dei testi assegnati.

In sede di conferenza saranno presentati i risultati ottenuti dalla compilazione dei questionari, che verranno discussi sia in relazione alle aspettative qui succintamente presentate, sia in ottica più generale, ponendo l'accento su fenomeni particolari eventualmente evidenziati dalle risposte ottenute.

Al di là della valutazione specifica dei singoli laboratori, può essere interessante una riflessione generale sul modello integrato ricerca-didattica proposto: se da un lato è innegabile il vantaggio di portare gli studenti a impegnarsi in un lavoro concreto, inserito in un contesto di ricerca reale e attivo, del quale apprezzare la multidisciplinarietà degli approcci e delle competenze in gioco, dall'altro si corre il rischio di concentrarsi più sull'obiettivo pratico che sulla reale acquisizione di competenze spendibili in altri contesti, come evidenzia in parte la strategia didattica messa in campo nel primo laboratorio.

Un'ulteriore criticità da non sottovalutare è la possibilità di perdere di vista il contesto didattico in cui la proposta si inserisce; ragione per cui nel secondo caso è stato maggiormente valorizzato anche l'aspetto filologico. Da questo punto di vista, il secondo esperimento sembra più equilibrato nel coniugare le necessità concrete del progetto di ricerca a quelle didattiche, ma ciò che certamente rende entrambi i laboratori un ottimo esempio di inserimento delle DH in un corso di studi tradizionale è la possibilità data agli studenti di toccare con mano le ricadute indirette della marcatura sullo studio del testo: cimentarsi in operazioni di codifica ha reso possibile una maggior consapevolezza delle problematiche testuali da affrontare, oltre che una diretta pratica della filologia d'autore, soprattutto nell'emendazione di errori tramandati dalle edizioni collazionate e nella conservazione di forme d'autore. Infine, l'attenzione riservata alla marcatura delle opere citate dall'autore ha permesso una riflessione sull'intreccio tra testo e fonte, essenziale per comprendere a pieno non solo la singola opera, ma l'intero impianto teorico e metodologico sotteso, il pensiero dell'autore colto nel dialogo con i propri modelli.

**RINGRAZIAMENTI**

Desidero concludere il mio intervento con un doveroso ringraziamento a chi ha progettato e reso possibile questo interessante esperimento didattico: la professoressa Paola Italia, che ha

fortemente voluto il laboratorio, oltre ad averlo concretamente coordinato e introdotto, e la professoressa Francesca Tomasi, che mi ha fornito un paziente e fondamentale supporto tecnico e metodologico.

## **BIBLIOGRAFIA**

- [1] Italia, P., Tomasi, F. 2014. *Filologia digitale. Fra teoria, metodologia e tecnica*. In: *Ecdotica*, 11, pp. 112-131.
- [2] Dariah teach <https://teach-blog.dariah.eu>
- [3] Tei by Example <http://teibyexample.org>
- [4] Pace R., 2015. *Digital humanities, una prospettiva didattica*, Roma, Carocci. ISBN: 9788843079575
- [5] Hockey, S. 2001. *Electronic Text in the Humanities: Principles and Practice*, Oxford, Oxford University Press.
- [6] Orlandi T., 2010. *Infomatica testuale. Teoria e prassi*. Laterza.
- [7] Fiormonte, D., 2008. Il testo digitale: traduzione, codifica, modelli culturali, in *Italianisti in Spagna, ispanisti in Italia: la traduzione. Atti del Convegno Internazionale (Roma, 30 - 31 ottobre 2007)*, Roma, Kappa, 2008, pp. 271-284.

# Storia dell'informatica: metodi e strumenti per raccontarla da informatici

Viola Bongini, Giovanni Antonio Cignoni, Emanuele Lenzi, Nicolò Fratelli  
HMR Project, Italy, giovanni.cignoni@progettohmr.it

## SOMMARIO

*Hackerando la Macchina Ridotta* (HMR) è un progetto di ricerca in storia dell'informatica che, accanto ai tradizionali metodi di indagine storica, usa molta informatica: competenze, per capire bene ciò che si studia, e tecnologie, per esempio per ricostruire con la simulazione software l'hardware del passato. HMR racconta anche la storia dell'informatica cercando di coinvolgere il grande pubblico altrimenti esposto alle vulgate che mostrano una disciplina scientifica come un hobby per nerd. Anche in questa missione è contemplato l'uso dell'informatica.

L'articolo presenta in particolare le iniziative per razionalizzare la presenza di HMR su internet e l'uso di web e social network come ulteriori canali per perseguire una comunicazione corretta della storia dell'informatica. Sono discussi tre lavori di tirocinio e tesi di studenti del corso di Storia dell'Informatica della laurea in Informatica Umanistica dell'Università di Pisa. Le esperienze, oltre alla componente tecnologica, hanno affrontato anche i contenuti, o meglio lo stile con cui comunicarli. Anche in questo caso si è attinto all'informatica ispirandosi alle prassi di scrittura controllata che hanno le loro radici nella tradizione del software requirement engineering.

## ABSTRACT

### **Computer science methods and tools for storytelling the history of computing**

*Hackerando la Macchina Ridotta* (Hacking the Smaller Machine, HMR) is a research project about history of computing. Alongside the methods of historical investigation, HMR uses computer science: technical knowledge to deeply understand old hardware and software, as well as technologies like software simulation to rebuild old machines. Moreover, HMR is committed to telling a correct history of computing, trying to oppose the stories that depict computer science as a hobby for nerds. In this mission too, information technologies are, of course, applied.

The article presents the work done to rationalise the presence of HMR on the internet and the use of web and social networks as additional channels to disseminate a correct version of the history of computing. The work stems from three final assignments of students of the course in History of Computing belonging to the Digital Humanities degree of the University of Pisa. These experiences, besides the technological component, addressed also the content or, better, the style needed to communicate it. In this case too, computer science is involved: guidelines were defined by adopting the controlled writing practices that have their roots in software requirement engineering.

## PAROLE CHIAVE (KEYWORDS)

computing history, communication, teaching, web, social network.

## INTRODUZIONE

*Hackerando la Macchina Ridotta* (HMR) è un progetto di ricerca in storia dell'informatica. Attivo dal 2006, affronta la storia dell'informatica usando soprattutto... l'informatica. In

primo luogo come competenze: coltivandole e usandole con maniacale attenzione per ottenere piena comprensione tecnica dei documenti e dei manufatti del passato. Poi adottando metodi e tecnologie dell'informatica di oggi a supporto dell'indagine storica: dalla simulazione software, al pattern matching sulle vecchie ROM. I risultati hanno premiato l'approccio: dalla riscoperta e ricostruzione del primo calcolatore italiano [1, 2], ai giochi educativi [3], alle recenti indagini sui PC oltrecortina [4]. Il nome del progetto ricorda la curiosità ossessiva degli hacker originali e il nomignolo della prima Calcolatrice Elettronica Pisana, oggetto degli inizi di HMR.

Comunicare i risultati della ricerca di HMR a un pubblico più ampio possibile è un obiettivo motivato dall'evidenza che l'*informatica percepita* è molto distante da quella reale, a partire dalla sua storia. Convenienze di mercato amplificate dai media sempre in cerca della notizia sensazionale presentano un'informatica fatta di gadget hardware e software, che ha una storia breve costruita sugli artifici retorici della semplificazione e dell'esagerazione, concentrata sui successi di pochi imprenditori o sulle vicende degli scienziati le cui biografie offrono materia facile a narratori e sceneggiatori – e.g. Turing, eroe della guerra dei codici con una drammatica storia personale. A tutto ciò si sovrappone una crescente disinvoltura nell'applicare le tecniche di storytelling alla comunicazione scientifica [5].

Lo scostamento c'è anche fra gli addetti ai lavori che, per dire, devono riflettere un po' prima di riconoscere che i contributi al "digitale" di Fibonacci e Pascal sono più rilevanti di quelli di Jobs e Gates. Colmarlo, per l'informatico-informatico, è un fatto personale, di conoscenza matura della propria disciplina, per l'informatico-umanista può essere una missione. In questo senso, la diffusione dei propri risultati e il contributo a una narrazione corretta della storia dell'informatica sono diventati per HMR occasione di sperimentazione. A percorsi tradizionali (mostre, didattica museale, eventi), sono state affiancate esperienze diverse come la collaborazione con un giornale online [6] o la realizzazione di allestimenti suggestivi [7].

L'articolo presenta tre iniziative recenti per razionalizzare la presenza di HMR su internet usando web e social network come canali aggiuntivi per la missione di comunicazione del progetto. I risultati sono stati ottenuti coinvolgendo in tre coppie tirocinio-tesi (due concluse, una prossima al traguardo) altrettanti studenti del corso di Storia dell'Informatica della laurea in Informatica Umanistica dell'Università di Pisa. Le tre esperienze, descritte nelle sezioni seguenti, oltre alla componente tecnologica, hanno affrontato anche i contenuti, o meglio le modalità con cui comunicarli: di nuovo si è attinto all'informatica riprendendo e adattando alcune prassi di scrittura controllata che hanno le loro radici nel *software requirement engineering*.

## **EPICAC, L'INFRASTRUTTURA WEB DI HMR**

HMR ha un sito web sin dal 2009, inizialmente ospitato dal Dipartimento di Informatica dell'Università di Pisa, usato per pubblicare le copie digitali dei documenti di archivio, il software prodotto, le presentazioni usate in seminari ed eventi, gli articoli scientifici. Per avere maggior libertà di sviluppare applicazioni web e per dare agli studenti coinvolti completo accesso a file e macchine, nel 2017 è stato registrato il dominio *ProgettoHMR.it* con server e piattaforma applicativa indipendenti.

Il primo passo della riorganizzazione della presenza internet di HMR è stato realizzare un semplice Content Management System (CMS) specializzato per le esigenze di archiviazione e comunicazione del progetto. Affrontato in un tirocinio curricolare e successiva tesi [8] e preceduto da una riflessione sui temi della comunicazione dei progetti di ricerca, il lavoro ha rivisto la struttura del sito web di HMR, ne ha rinfrescato la grafica, ha realizzato una piattaforma CMS dedicata e, infine, ha traslocato tutti i contenuti del vecchio sito a febbraio 2018.



La decisione, impegnativa, di realizzare un CMS dedicato è stata presa dopo un'attenta analisi di pro e contra. In breve: le applicazioni disponibili non soddisfacevano tutte le esigenze del progetto mentre sviluppare o modificare estensioni di soluzioni esistenti (e.g. plug-in di *Wordpress*) richiedeva competenze specialistiche e introduceva dipendenze tecnologiche. Il CMS realizzato, *EPICAC* (Easy Project Information and Content Administration and Communication, omaggio a un calcolatore protagonista di un racconto di fantascienza [9]), oltre a gestire contenuti e accessi, fornisce l'infrastruttura per integrare altre applicazioni web (vedi sez. successive). La maggior parte dei contenuti web di HMR sono oggi gestiti via *EPICAC* (oltre 400 pdf e 250 immagini al tempo della migrazione). Per scelta è stata mantenuta la libertà di poter realizzare pagine HTML manualmente, ma sono stati previsti elementi CSS e funzioni JS per semplificare al massimo il lavoro del redattore.

## **OGGI STI, ACCADDE OGGI NELLA STORIA DELL'INFORMATICA**

Interessare il pubblico su argomenti di storia dell'informatica senza ricorrere alle rivisitazioni della finzione narrativa o cinematografica e, purtroppo, anche di molta saggistica di cassetta, è una sfida. OggiSTI la raccoglie raccontando la storia dell'informatica per eventi quotidiani sfruttando la curiosità per ciò che "accadde oggi". Il fuoco sul singolo giorno permette di dare sufficienti dettagli mantenendo brevi i testi e concentrata l'attenzione del lettore.

OggiSTI è anche un esperimento di narrazione collaborativa. L'applicazione gestisce il workflow per il popolamento dei contenuti da parte di utenti autenticati e organizzati in redattori e revisori. La redazione degli eventi è aperta a tutti, ma prima della pubblicazione i testi necessitano dell'approvazione dei revisori. Nella realtà del web dove molti contenuti non sono controllati o sono affidati al parere della maggioranza, alle azioni delle lobby più attive o al verdetto dei "like", OggiSTI reintroduce la *peer review* preventiva tipica delle comunità di ricerca. Il modello che ne emerge – redazione aperta e pubblicazione verificata – sembra una buona soluzione per avere i benefici di una partecipazione ampia senza i rischi delle correzioni a posteriori (se e quando) tipiche di *Wikipedia*.

L'infrastruttura applicativa di OggiSTI è stata sviluppata in un tirocinio curriculare seguito da tesi [10]. Alla redazione dei contenuti partecipano come progetto didattico gli studenti del corso di Storia dell'Informatica. I progetti impegnano gli studenti su piccole ricerche storiche condotte con attenzione soprattutto alle fonti, quasi sempre dimenticate nel caos della rete. Segue l'esercizio di scrittura, definito negli spazi e nello stile che non ammette toni pubblicitari, aggettivazione gratuita, approssimazione tecnica: le linee guida per i redattori di OggiSTI stabiliscono vincoli che somigliano molto alle regole di scrittura dei requisiti software espressi in linguaggio naturale [11, 12].

## **FACEBOOK, UN PROGETTO EDITORIALE**

HMR ha atteso molto prima di approdare ai social network. Le potenzialità erano note, tuttavia, anche alla luce di visioni che, passati gli entusiasmi iniziali, guardano al fenomeno con maggior spirito critico [13], non si è voluto buttarsi senza aver predisposto dei criteri che permettessero di usare gli strumenti con efficacia ed efficienza. La prima riguarda la missione di comunicare correttamente la storia dell'informatica. La seconda è necessaria considerando le limitate risorse disponibili all'interno del progetto – "postare" bene richiede lavoro.

L'occasione, per adesso limitata al solo Facebook, si è presentata con la solita coppia tirocinio curriculare più tesi che sono tuttora in corso. La componente tecnologica riguarda l'integrazione fra OggiSTI e FB: parte dei contenuti pubblicati per ricordare l'evento quotidiano sono condivisi automaticamente sulla pagina FB, il post è breve ed essenziale, chi è incuriosito può continuare la lettura sulla pagina web.

I post automatici alleviano parte del lavoro, ma un'attività redazionale "manuale" rimane e la tesi ha sviluppato le linee guida per la scelta dei contenuti e lo stile del linguaggio. Anche in questo caso, riprendendo tecniche nate nell'ambito dell'ingegneria del software, la libertà di espressione del linguaggio naturale è controllata per evitare incomprensioni, ambiguità e imprecisioni; il redattore ha indicazioni chiare, il lettore trova testi più facili da leggere.

Sono infine stati impostati dei piani editoriali come, per esempio, il post domenicale che condivide una vignetta a tema informatico commentandola in prospettiva storica, oppure le modalità per annunciare e promuovere le iniziative di HMR o gli eventi a cui HMR partecipa. Nel lavoro di tesi è previsto anche un periodo di osservazione per misurare i risultati, sia come indicatori relativi a FB sia come ricadute sulle visite alla pagina web di HMR.

## **RISULTATI (PREMATURI) E RIFLESSIONI**

HMR è un progetto più che decennale, ma la razionalizzazione della sua presenza internet è un lavoro in corso: le iniziative descritte sono recenti ed è presto per apprezzare grandi cambiamenti. EPICAC è un miglioramento infrastrutturale, OggiSTI copre ancora solo un terzo dei giorni dell'anno, la comunità FB è appena agli inizi.

Gli accessi alle pagine web sono ancora molto legati al corso. Con FB sta aumentando l'insieme di persone a conoscenza di HMR; i numeri sono già interessanti (circa 2k e 1k come medie settimanali di copertura e interazione) e in crescita, ma non c'è l'effetto sperato sugli accessi alle pagine web; a FB è imputabile un +15% nell'ultimo semestre: i post "girano" ma pochi poi approfondiscono. Nella storia di HMR rimane imbattuto il periodo 2014/15 quando il pubblico "esterno" arrivava sulle pagine di HMR grazie agli articoli seriali pubblicati su un giornale online [6]: la chiusura della testata ha lasciato un vuoto che, per ora, FB non ha colmato. Fra gli utenti di FB probabilmente la percentuale di quelli portati a saperne di più è, fisiologicamente, più bassa rispetto ai lettori di un giornale.

Come prima analisi si possono azzardare due linee d'azione complementari: da una parte accettare la percentuale bassa di "approfondenti" ma contare sui numeri che FB può raggiungere, dall'altra proporre al resto dei "non approfondenti" pillole di buona storia che comunque aiutino ad avvicinare l'informatica percepita a quella reale.

## **RIFERIMENTI**

- [1] Cignoni, G.A., Imbrenda C. e Ceccarelli D. 2009. Il "restauro" del software di sistema della Macchina Ridotta del 1956. Atti del 47mo Congresso Nazionale AICA. AICA.
- [2] Cignoni, G.A. e Gadducci F. 2012. Rediscovering the Very First Italian Digital Computer. Atti di 3rd History of Electro-technology Conference, HISTELCON 2012. IEEE.
- [3] Cignoni, G.A., Mongelli T. e Cappellini L. 2015. Games, from Engaging to Understanding: a Perspective from a Museum of Computing Machinery. Atti della 14th International Conference on Entertainment Computing. IFIP.
- [4] Bodrato S., Caruso F. e Cignoni G.A. 2018. Discovering Eastern European PCs by hacking them. Today. Apparirà negli atti di *Histories of Computing in Eastern Europe*. IFIP.
- [5] Katz Y. 2013. Against storytelling of scientific results. In *Nature Methods* 10: p. 1045
- [6] Cignoni, G.A. e Colosimo C. 2017. Raccontare il calcolo senza fare i conti. In *Museologia Scientifica Memorie*, n. 16/2017.
- [7] Cignoni, G.A. e Magnani. A. 2017. Prima dell'emmepitré pixel a 45 giri: contatti fra musica e tecnologia. Intervento a *Il museo e i suoi contatti - XXVII Congresso dell'Associazione Nazionale dei Musei Scientifici*, Genova, 25-27 ottobre 2017.

- [8] Lenzi, E. 2018. Comunicazione e progetti di ricerca: una soluzione CMS per HMR. Relazione di laurea in Informatica Umanistica (relatori G.A. Cignoni, S. Turbanti).Università di Pisa.
- [9] Vonnegut, K. 1950. *EPICAC*. In: *Collier's Weekly*. Novembre 25, 1950.
- [10] Pratelli, N. 2017. Un'applicazione web: Oggi nella storia dell'informatica. Relazione di laurea in Informatica umanistica (relatori G.A. Cignoni, E. Salvatori). Università di Pisa.
- [11] **Fuchs**, Norbert E (a cura di). 2009. Controlled Natural Languages. Lecture Notes in Computer Science vol. 5972, Springer.
- [12] Cheng, B. e J. Atlee. 2007. Research directions in requirements engineering. Atti di *Future of Software Engineering FOSE07*, IEEE.
- [13] Martinez, A.G. 2016. Chaos Monkeys: Obscene Fortune and Random Failure in Silicon Valley. HarperCollins.

# From collaborative transcription to interdisciplinary education: the postcards of the Great War case

Enrica Salvatori\*, Federico Boschetti°, Angelo Mario Del Grosso°

\*Laboratorio di Cultura Digitale, Università di Pisa, enrica.salvatori@unipi.it;

°Istituto di Linguistica Computazionale “A. Zampolli” (ILC-CNR), Area della ricerca di Pisa, {nome.cognome}@ilc.cnr.it

## ABSTRACT

The authors intend to present an educational experimentation that took place in the first semester of the academic year 2018-2019 at the degree course in Informatica Umanistica at the University of Pisa. The experimentation involved the courses of Digital Public History, Digital Text Encoding as well as Digital Philology, and concerned the digitization of a corpus of postcards from the period of the First World War.

The postcards, owned by the ethnographic museum of La Spezia “G. Podenzana” and in collaboration with it, have been historically contextualized, digitized, placed on a platform and distributed to the students in order to be recorded, transcribed and encoded with the following purposes:

- make students participate in a real Digital Public History project;
- consider the problems of how to record, transcribe, share and encode a corpus of historical postcards;
- evaluate how to combine more educational activities into a wider project of crowdsourcing;
- test a collaborative platform and some encoding tools;
- test methods and timing to allow the collaboration between individual courses within a university degree course in DH.

The authors aim to highlight the main problems encountered and the added values of the integration of courses in teaching Digital Humanities.

## KEYWORDS

Digital Public History, Collaborative Philology, Text Encoding, Digital Philology, Web Application, Educational, Digital Textual Scholarship

## INTRODUCTION

Recording, digitizing, encoding and publishing postcards are suitable activities for educational and training purposes in the context of Digital Humanities classes. Indeed, postcards are quite complex documents, due to their different main description levels, textual divisions and modalities of communication.

Specifically, postcards consist of images, printed texts, short handwritten texts, structured data and formatted information arranged within the destination form and stamps. From the catalographic point of view, for example, they show a dual nature, bibliographic (printed in series by a publishing house/typography with images and writing) and archival (manuscript), without counting all that concerns the composite area of philately. Moreover, the handwritten message of a postcard generally follows a well structured sequence: opening, body and closing. The digital encoding of metadata for this kind of documents is very rich and must

take into account location, dimensions, subjects, materials, preservation status, etc. During the digitization process the image which appears on the verso of the card as well as the stamps on the recto (postmark and postage) must be properly described.

Many initiatives aim at transcribing, studying and publishing archives of postcards; among the others, the CARLI Digital Collections, and *Transcribathon* into Europeana or some parts of the DALF project are noteworthy. However, our initiative is different because we mainly focus our attention to the engagement of DH students in building the digital archive and mastering all the phases of the activity in order to prepare and test a project addressed to a wider public.

## METHODOLOGY

In the Digital Public History course, the context of production and use of postcards - in addition to the peculiarities related to their recording - have been explained. In order to understand the aims and the methods of a crowdsourcing project, specific stakeholders have been invited to participate. School class assignments were also distributed in accordance with Digital Text Encoding and Digital Philology teachers.

As far as the Digital Text Encoding course is concerned, the exploitation of postcards has been extremely useful to engage students with a real project full of challenging issues. Each student received six postcards to encode in order to apply their skills to real world cases. The encoding job has been carried out by each student in collaboration with their course companions. The TEI guidelines<sup>6</sup> define a set of best practices and elements to suitably encode all the textual phenomena which occur in the original materials, thus it has been possible to explain and practice with different TEI modules conceived to work, amongst other, with manuscripts and named entities.

Concerning the Digital Philology course, students have been involved in the extension of Euporia, a tool created at the CNR-ILC for the annotation of primary and secondary sources by XML-TEI or by Domain Specific Language. The new module for Euporia takes into account the specific layout for the annotation of postcards (as illustrated in Fig. 1), constituted by a header, more the first and the second side of the postcard with the related TEI encodings. The annotation editor, based on the well-known ACE Editor, is flexible enough to support different grammars to validate the annotation or the encoding.

---

<sup>6</sup> <http://www.tei-c.org/guidelines/>



Fig. 1 – Euporia Collaborative Web Application for text encoding: Artistic postcard, Women

## DISCUSSION

The organizational context of the Italian university does not facilitate the integration between individual classes even within a well built degree course as that one of Informatica Umanistica in Pisa: problems of agenda and fixed hours of lesson have proven themselves concrete obstacles as well as the standard shape and furnishing of the classrooms, that usually do not help group work. The organization of work would need a stable help (tutor), able to deal with all the three courses involved, but the division of the tasks, dosed according to the interests of the people and the availability of trainees and undergraduates, has overcome the greater problems encountered. Other solutions, that we could define in a certain sense extemporaneous but sometimes effective, have involved the cross participation of teachers within the colleagues' courses (although not easy to organize due to schedule issues) or the use of some well-known instant messaging applications (such as whatsapp or skype) to solve specific problems arisen during the classroom discussion (as illustrated in Fig. 2).



Fig. 2 – Whatsapp real time chat session during a lesson.

Finally, at the end of the first part of the courses, students were asked to evaluate the operative part of the course itself, but not through an anonymous survey: we decide to make them think all together (in the Digital Public History course) around the matters they encountered and the possible solutions (including the potential load of work that the solutions would entail) and which changes to make in order to put in place a real crowdsourcing project. The outcome has been twofold: firstly, it has been collected a few suggestions to better organize future courses and, secondly, the problem solving teamwork that has been realized, has been actually useful both to understand the individual topic of interest, and also to be aware of the overall issues involved in the management of a DH project. A further evaluation on the courses interest will be available after the institutional survey carried out by each student before their exams. In any case it was clear to all the teachers involved in this initiative that the combination of courses within a common project gave professors and students a greater motivation to work and study in a more effective way.

The experimentation of integrating three courses of the same institution allows us to handle and show to the students all the possible phases involved in the process of digitizing and publishing historical textual materials. Moreover, our intercourses work allows students to exploit the skills acquired during different contexts and to profitably integrate them in one common project. In particular, while digital philology students improve their software design and software developing skills, they provide also the basic tools and the collaborative platform to facilitate the other students in encoding and saving the documents.

## CONCLUSION AND FURTHER WORKS

Since the initiative is still underway, the authors do not present any conclusions but they think only to highlight the main problems encountered and the added values of the integration of courses in teaching DH. Two of the three courses involved have been completed, the third will be concluded in the second semester.

Preliminarily, we can certainly say that the problem of lesson schedules could be resolved through a good organization among teachers during the educational planning phase and then by keeping common slots during the construction of the timetable. These joint lessons are not easy to obtain, from one hand because they do not occupy periodic and stable slots and, from the other hand, because they need specific classroom posing in this way some problems in the classroom management system. For this kind of integrated education, the university institution should think to specifically designated classrooms for workgroups within open spaces equipping them properly.

The problem concerning the different skills among students from different courses, specifically with regards to Digital Public History and Digital Text Encoding could be overcome, as already mentioned, by hiring tutors. Tutors have not to be necessarily external ones (in fact, already competent students recruited on a voluntary basis or as trainee or those being into graduation projects can be a satisfactory solution). Moreover, students claimed to set up guidelines before the start of courses, requiring in this way a lower level of coding details.

With regard to the actual work of transcription and recording, also in relation with the cultural institutions involved, the main problems have been found in the granularity of the data requested by the institutions (average), acceptable in a crowdsourcing project (low) and to be used in a university course (high).

It is the hope of the authors to present at the next AIUCD conference (2020) the main conclusions of the work by defining better the hurdles encountered and the solutions put in practice to overcome them.

## ACKNOWLEDGMENTS

Thanks to Giacomo Paolicchi, conservator of the "Giovanni Podenzana" Ethnographic Civic Museum and Marzia Ratti, director of museums and cultural services, Municipality of La Spezia.

## REFERENCES

- [1] Booth A. 1996, *Postcards from the Trenches: Negotiating the Space Between Modernism and the First World War*, Oxford University Press.
- [2] Burnard L. 2014, *What is the Text Encoding Initiative?* Encyclopédie numérique 3. OpenEdition Press.
- [3] Cati I. 2006, *Cara mamma ti scrivo: le cartoline dei soldati della grande guerra*, Gaspari.
- [4] Cole K.J. 2016, *Postcards from the Front 1914-1919*, Amberley Publishing Limited.
- [5] Delle Cave L. 2013, *Orme di guerra : lettere e cartoline dal fronte (1912-1919)*, Sarnus.
- [6] Del Grosso A. M., Cristofaro S., De Luca M.R., Giovannetti E., Marchi S., Seminara G., Spampinato D. 2018, *Le lettere di Bellini: dalla carta al Web*. In Associazione per l'Informatica Umanistica e la Cultura Digitale (AIUCD) 2018, 60–64. Bari: Umanistica Digitale.
- [7] Doyle P. 2011, *British Postcards of the First World War*, Bloomsbury Publishing.
- [8] Fochesato W. 2015, *Il gioco della guerra: l'infanzia nelle cartoline del primo conflitto mondiale*, Interlinea.
- [9] Mugelli G., Boschetti F., Del Gratta R., Del Grosso A.M., Khan F. and Taddei A. 2016, *A User-Centred Design to Annotate Ritual Facts in Ancient Greek Tragedies*, BICS 59 (2), 103-120.
- [10] Pierazzo E. 2015, *Digital Scholarly Editing : Theories, Models and Methods*. Digital Research in the Arts and Humanities. Farnham Surrey: Ashgate.
- [11] Pozzi, 2015, *Lettere dal fronte Poste Italiane nella Grande guerra*, Giuntini A. (éd.), Milano, Rizzoli.
- [12] Rosselli Del Turco R. and Di Pietro C. 2016, *Between innovation and conservation: the narrow path of UI design for the DSE*. In *Digital Scholarly Editions as Interfaces*. Graz: University of Graz.
- [13] Salvatori E. 2017, *Digital (Public) History: la nuova strada di una antica disciplina*, RiMe, 1-I, 57-94.



# Mapping the Persecution of Trentino People Deported to the 3rd Reich Camps

Rachele Sprugnoli, Alessio Palermo Aprosio, Giovanni Moretti, Sara Tonelli

Digital Humanities Group, Fondazione Bruno Kessler, Trento  
*rachele.sprugnoli@alumni.unitn.it; [aprosio,moretti,satonelli]@fbk.eu*

## ABSTRACT

In this proposal we describe the results of a project aiming at tracing the movements of Trentino people that were deported to the 3rd Reich camps during World War II. More specifically, we performed the semantic annotation, georeferencing and visualization of data collected by expert historians. This work wants to shed light on the stories of people, civilians but also military internees, that were forcibly moved to labour, detention or extermination camps and, in this way, preserving and passing on their memory through the use of digital tools.

## KEYWORDS

Spatial humanities, digital history, RDF, WW2.

## INTRODUCTION

This abstract presents a project whose goal is to map the persecution of 202 persons of Trentino that were deported to the 3rd Reich camps during the Second World War and give access to the corresponding dataset both in RDF and through an interactive visualization. This work takes advantage of an online archive<sup>7</sup> made available by “Laboratorio di storia di Rovereto” whose data have been published also in a dedicated volume [1].

The processed data contain information about victims who were born in the Trentino province but also people born in the municipalities passed to South Tyrol after the De Gasperi-Gruber agreement<sup>8</sup> and emigrants whose family was originally from Trentino. Among these people there are different types of deportees: opponents of Nazism, emigrants captured in France, civilians working in Germany, veterans of the Spanish civil war, military internees and prisoners, draft evaders and deserters, civilians convicted for political reasons. Although each of these people had a different story of persecution, their fate is in common, marked by forced displacements and inhumane experiences that have brought most of them to death.

## RELATED WORKS

Nazism was based on a strong spatial theory [4] that led to the territorial expansion of the Third Reich, the deterritorialization of masses of people forcibly moved from their native context, and the creation of new places (e.g. forced labor, transit and extermination camps). The geographical dimension of the Nazis’ persecutory strategy has in particular been studied by taking into consideration the various spatial processes that have characterized it, e.g. deportation and concentration. Different approaches have been adopted to study these processes as exemplified by the books “Hitler’s Geographies” [5] and “Geographies of the Holocaust” [6]. In the former, Stone [7] highlights two important points that constitute our theoretical basis.

---

<sup>7</sup> [http://www.labstoriarovereto.it/archivi/trentini\\_deportati\\_nei\\_Lager\\_del\\_3\\_Reich](http://www.labstoriarovereto.it/archivi/trentini_deportati_nei_Lager_del_3_Reich)

<sup>8</sup> [http://www.provincia.bz.it/pariservertrag/index\\_i.asp](http://www.provincia.bz.it/pariservertrag/index_i.asp)

First of all, Nazi persecution affected the whole continent: it was not just a matter related to specific and sadly famous sites such as Auschwitz or the Warsaw ghetto but histories of violence have happened everywhere. Our focus on Trentino is placed in this perspective, shedding light on episodes of persecution that destroyed families and local communities. Secondly, Stone notes that victims' testimonies are to be considered as “spatial trajectories” [8][9]: the narration of persecutions is rooted in the places where these people were dislocated and segregated and made by sequences of movements between places. Our visualization wants to make these trajectories evident plotting them on a map and linking them also to a timeline. In this way, both individual experiences and general patterns can be identified. Our geovisualization approach to study the geographical dimension of Nazis genocidal policy is in line with previous efforts [10] [11] but we use an original dataset whose release in RDF can benefit the research community and give visibility to a piece of Trentino and Italian history.

## WORKFLOW

Given that the needed information were not stored in a structured format, we performed the semantic annotation of each web page of the online archive using PundIt Annotator Pro<sup>9</sup> Chrome extension [2] personalized with the predicates of the Shoah ontology [3]. Indeed, although the victims of Trentino considered in this study were not Jewish, their persecution followed the same phases formalized in that ontology: e.g. arrest, detention, deportation to a Nazi camp, transfer to other camps. A screenshot of PundIt annotations is displayed in Figure 1<sup>10</sup>. As in the data on the Italian victims of the Shoah released by Centro di Documentazione Ebraica Contemporanea<sup>11</sup>, details about the persecution are connected to biographical information, such as the date and place of birth.

---

<sup>9</sup> <http://thepund.it/annotator-pro-web-annotation/>

<sup>10</sup> Part of this annotation was done by a high school student during his internship in the context of a work-school alternation project.

<sup>11</sup> <http://dati.cdec.it/lod/shoah/reference-document.html>

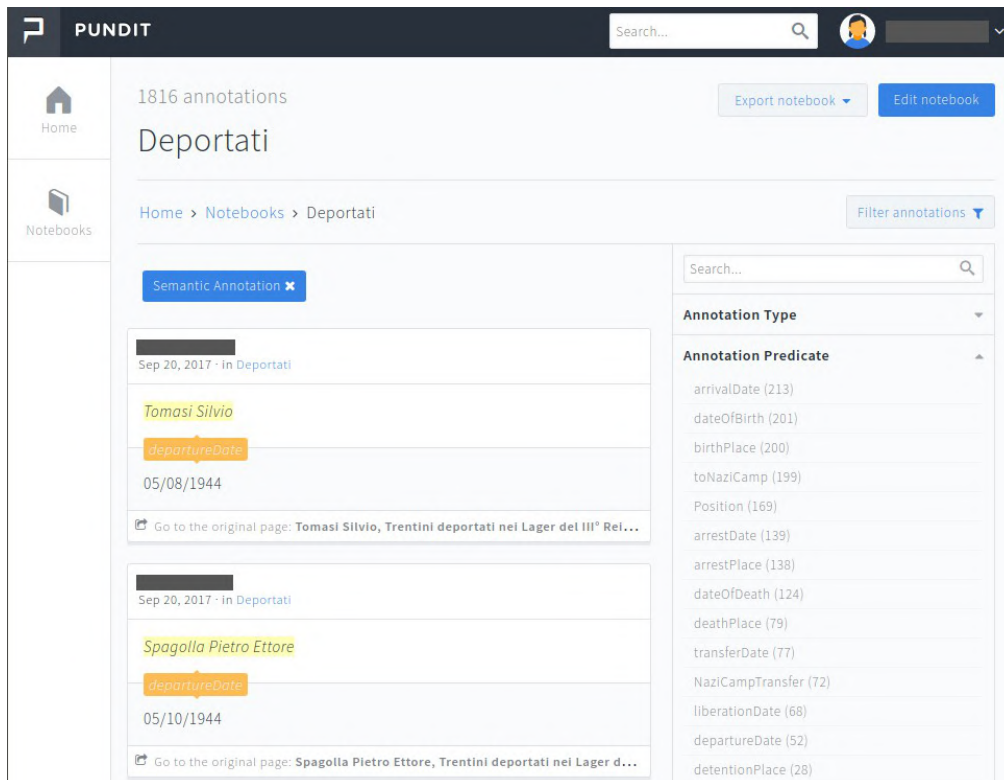


Figure 1. Screenshot of the semantic annotations of the online archive with the list of predicates taken from the Shoah ontology.

The output of the semantic annotation was then converted in JSON, places mentioned in the dataset were georeferenced using Nominatim<sup>12</sup> and the story of the victims reconstructed through a visualization showing the trajectories between the different stages of the persecution [12]. Finally, RDF triples were parsed and released in Turtle format<sup>13</sup>. Both the RDF data and the stand-alone visualization application are available for download<sup>14</sup>.

## DATA ANALYSIS AND VISUALIZATION

Our dataset contains information about 202 persons (199 men and 3 women): the strong majority of them (>90%) were not born in Trento, but in almost one hundred small towns and villages scattered in valleys and mountains, such as Canal San Bovo and Castel Tesino. 32% of the victims arrived at Dachau and 21% at Mauthausen but different types of camps are registered as well, all belonging to the complex Nazi punitive system and all characterized by the same treatment for the victims: inhuman conditions of life, insufficient food and poor or no health care. In particular, there are workers' educational camps (e.g., Reichenau), prison labour camps (e.g. Bernau am Chiemsee) and punishment camps (e.g. Allach). Few victims remain in the camp of their first matriculation: they are usually moved from one camp to another, or between subcamps, often with repeated transfers and sometimes returning to the initial camp. For example, Figure 2<sup>15</sup> shows the RDF graph of the persecution of Silvio Galeaz, a farmer from Val di Non, arrested as Nazism opponent and deported to Dachau:

<sup>12</sup> <https://nominatim.openstreetmap.org/>

<sup>13</sup> <https://www.w3.org/TR/turtle/>

<sup>14</sup> <https://dh.fbk.eu/technologies/tn-navigator>

<sup>15</sup> We created Figure 2 with the ontology-visualization tool (<https://github.com/usc-isi-i2/ontology-visualization>) and Graphviz (<https://www.graphviz.org/>).

within 9 months he underwent several transfers, first to Buchenwald then to Dora-Mittelbau and Ohrdruf, where he died a month after his arrival. In total, 53% of deportees did not survive.

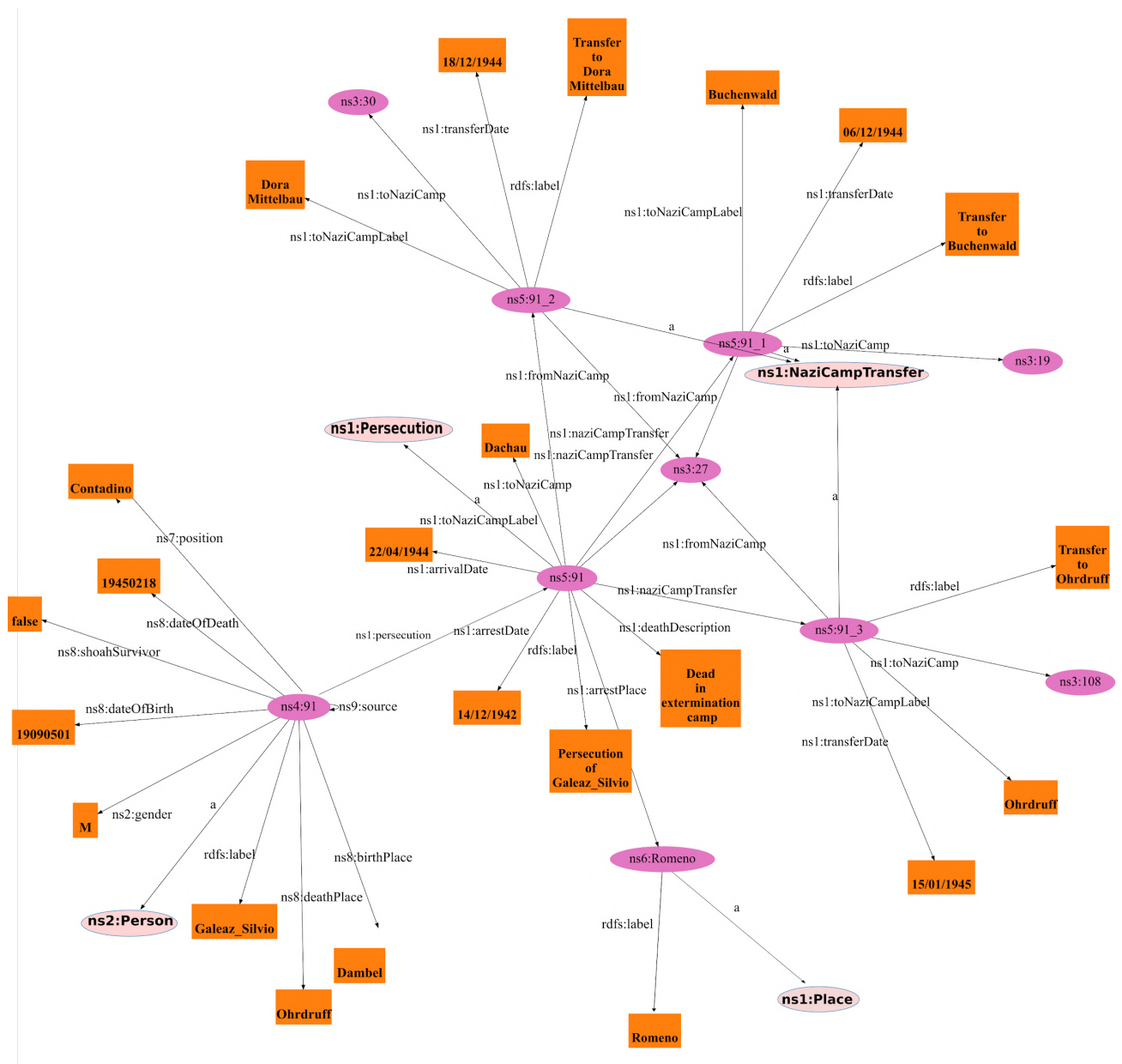


Figure 2. Visualization of the RDF graph representing the persecution of Silvio Galeaz.

All the trajectories can be navigated in an interactive stand-alone visualization application (see Figure 3) built using open-source framework: i.e., Node.js, Electron and Leaflet. Users can perform different types of searches, filter the results, follow a victim's story using a fly-over functionality and access the entry in the original online archive through a direct link.

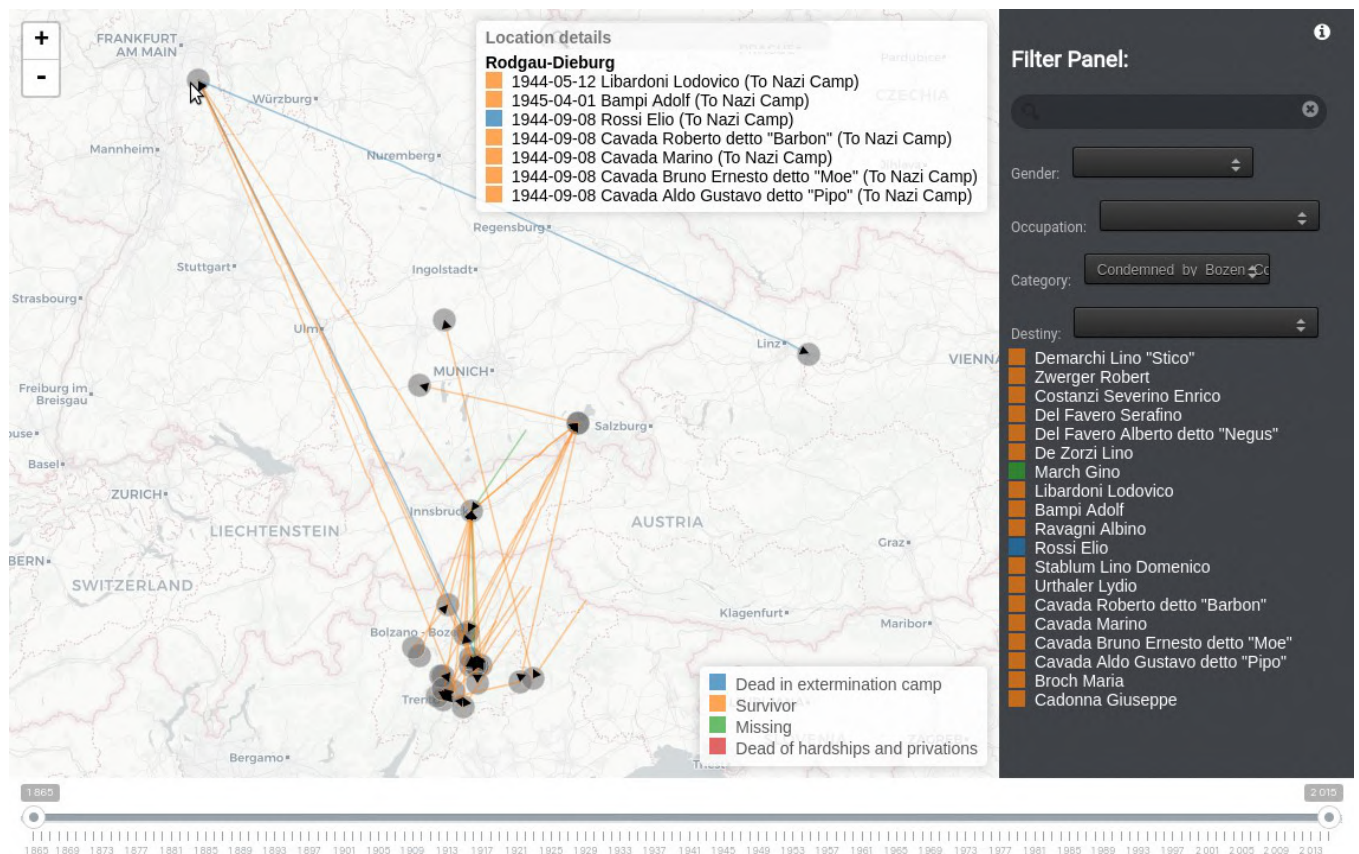


Figure 3. Persecution patterns of draft evaders and deserters condemned by the Bozen Court.

## ACKNOWLEDGMENTS

The authors want to thank Giulio Andreini and Francesca di Donato (Net7) for providing a personalized version of the PundIt Annotator Pro Chrome extension. We also thank Daniel Messina for performing part of the semantic annotation and Laboratorio Storico di Rovereto (in particular, Dr. Giovanni Tomazzoni) for publicly sharing the data that we used.

## BIBLIOGRAFIA

- [1] Laboratorio di storia di Rovereto. 2013. *Almeno i nomi. Civili trentini deportati nel Terzo Reich. 1939-1945*. Presidenza del Consiglio della Provincia autonoma di Trento. ISBN: 9788897372431.
- [2] Grassi, Marco, Christian Morbidoni, Michele Nucci, Simone Fonda, and Francesca Di Donato. 2013. "Pundit: Creating, Exploring and Consuming Semantic Annotations." In *Proceedings of the 3rd International Workshop on Semantic Digital Archives (SDA 2013)*, pp. 65-72.
- [3] Brazzo, Laura, and Silvia Mazzini. "From the Holocaust Victims Names to the Description of the Persecution of the European Jews in Nazi Years: the Linked Data Approach and a New Domain Ontology." *Book of abstract of DH (2015)*.
- [4] Barnes, T. J., & Minca, C. 2013. "Nazi spatial theory: The dark geographies of Carl Schmitt and Walter Christaller." *Annals of the Association of American Geographers*, 103(3), 669-687.
- [5] Giaccaria, Paolo, and Claudio Minca, eds. *Hitler's geographies: The spatialities of the Third Reich*. University of Chicago Press, 2016.

- [6] Knowles, Anne Kelly, Tim Cole, and Alberto Giordano, eds. *Geographies of the Holocaust*. Indiana University Press, 2014.
- [7] Stone, Dan. 2016. "Holocaust Spaces". In *Hitler's Geographies: The Spatialities of the Third Reich*, 45.
- [8] Gigliotti, Simone. 2017. *The train journey: Transit, captivity, and witnessing in the Holocaust*. Berghahn Books.
- [9] De Certeau, Michel.. 1984. *The practice of everyday life*. Berkeley.
- [10] Beorn, Waitman, Tim Cole, Simone Gigliotti, Alberto Giordano, Anna Holian, Paul B. Jaskot, Anne Kelly Knowles, Marc Masurovsky, and Erik B. Steiner. "Geographies of the Holocaust." *Geographical Review* 99, no. 4 (2009): 563-574.
- [11] Giordano, Alberto, and Anna Holian. "Retracing the "hunt for jews": A spatio-temporal analysis of arrests during the holocaust in Italy." In *Geographies of the Holocaust*. Indiana University Press, 2014.
- [12] Menini, Stefano, Rachele Sprugnoli, Giovanni Moretti, Enrico Bignotti, Sara Tonelli, and Bruno Lepri. 2017. "Ramble on: tracing movements of popular historical figures." In *Proceedings of the Software Demonstrations of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 77-80. 2017.

# Text in metamorphosis: new aesthetics of digital humanities

Herbert Natta

University of Rome - Tor Vergata, IT, herbert.natta@gmail.com

## ABSTRACT

The idea of *text* is etymologically founded on the linear and two-dimensional structure of the textile weave. However, the introduction of digital tools in humanities has progressively changed this metaphorical ontology, reshaping the configuration, representation and imagination of texts in different kind of forms. From Franco Moretti's *Graphs, maps, trees* to the linguistic treebanks, from the overproduction of literary atlases and cartographies to the more recent neural network language processing, the research in different fields, and in particular in digital humanities, has contributed to reimagine the iconic structure of texts. If, as Giuseppe Di Napoli observes, «our first cognitive act is to recognize the form of the things» (Di Napoli, 2011: XVI), and, with George Lakoff, «our ordinary conceptual system [...] is fundamentally metaphorical» (Lakoff, 2003: 4), this metamorphosis determines a relevant change in the way we perceive texts.

Starting from an historical overview that, from pre-digital to the contemporary post-human era, examines examples of this transformation, this work will propose, as a case study, a research on language visualization, where the composition of the iconic structure of the output aims at translating, in a relational space, the centre-periphery dichotomy.

## KEY-WORDS

aesthetic, visualization, metamorphosis, text, form

## BACKGROUND

The writing technology established the structure of texts as linear (Ong, 1982). The «method was to tear the elements of the image (pixels) from the surface and arrange them into lines» and, consequently, it has been «transcoded the circular time of magic into the linear time of history» (Flusser, 2000: 10).

The invention of printing closed this linearity in a recursive two-dimensional space, encouraging a sense «that what is found in a text has been finalized, has reached a state of completion», representing «the word of an author in definitive or 'final' form» (Ong, 1982: 129).

However, in parallel with the establishment of this standard, it is possible to find, in different ages, several examples of other textual configurations: a line of artistic exploration that has become more relevant, since the beginning of the XX<sup>th</sup> century, for the avant-garde movements (Dada, Futurism, Surrealism, etc.).

They progressively violate the calligraphic, orthographic, typographic restrictions, creating «new relationships between words and things» (Magritte, 1929: 32-33). The experience of the Avant-gardes has represented the theoretical and formal premise of the next experimental poetry, where it is possible to situate the beginning of digital literature.

Not accidentally, Max Bense, one of the main exponents of the post-war concrete poetry, found his ‘information aesthetic’ «on the grounds of modern science and technology» (Cramer, 2005: 3), using as a reference the information theory of Claude Shannon<sup>16</sup>.

In 1959, Theo Lutz, an IT engineer who cooperated with Bense at the University of Stuttgart, published on «Augemblick», a review of avant-garde poetry, *Stochastische Texte*, the result of an analysis, based on Shannon’s theory, of Franz Kafka’s novel *The castle* and its recombination through an algorithm developed from the Markov chain model.

This first experience of electronic literature revealed a radical transformation of the traditional relationship between writing, text and reading. The author writes the algorithm, the machine decomposes and recomposes the old text in a new one, the reader observes the performance: «the textuality of this electronic art may ‘thin out’ the complexities of reception or undercut the (constitutive) emphasis on active readership — on reading as much more than the sensation of being caught up in (or cannibalized by) an algorithm» (Andrews, 2003).

«Computers are better at poetry than they are at math», observed Jacques Leslie (Leslie, 1995), confirming the idea expressed by Theodor Nelson a decade before (Nelson, 1981).

However, the metamorphosis of texts in the digital era doesn’t concern just literary and poetic works: the idea of the text as a source of computational data has produced many new fields of investigation in digital humanities.

Franco Moretti, for example, has based his «distant reading» on «a process of deliberate reduction and abstraction» (Moretti, 2005: 1), that implies the use of digital tools to obtain statistical graphs and maps from textual data. Otherwhere, in computational linguistic, the semantic analysis, the natural language processing, the sentimental analysis produces networks, word clouds, treebanks, neural networks: different kind of visualization and conceptualization that refers to various metaphorical resources, other than the traditional textile field.

This digital reshaping generates two relevant questions: from one side, it radically changes the way of perceiving and conceptualizing the idea of text. Far away from the eternal God’s word, it goes also out from the boundary of the final, fixed form of printed books, representing a kind of fluid, flexible substance, more similar to organic, biological systems than to solid monuments. Its dynamic metamorphosis, in some way, contravenes the idea of the stability of the written word. From the other side, the multidimensional structure of this digital representations requires, also for disciplines traditionally more focused on textual analysis, to adopt methods and tools of image analysis, from fields like aesthetic and neuro-aesthetic.

## OBJECTIVES

This historical background, along with various examples of digital visualization of texts, has been the starting point for this research project, focused on the development of a dynamic network visualization of language data, from the macroscale of world statistical data to the microscale of textual relational space, with the objective to represent and translate the central-periphery dichotomy into an iconic structure.

---

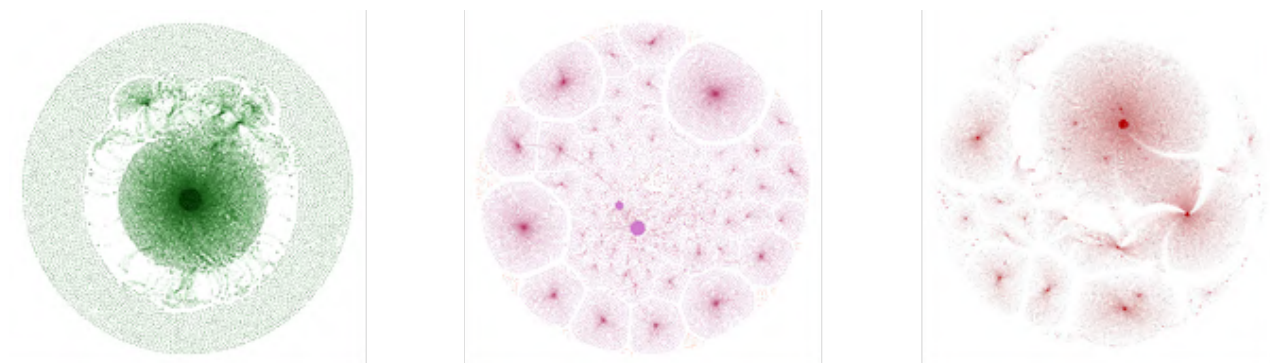
<sup>16</sup> In 1948, Claude Shannon, an engineer of the AT&T Bell Labs, formulated an information theory that distanced from classical semiotic, based on the relationship between signifier and meaning, in the direction of a quantitative information analysis. In *A mathematical theory of communication* (C. E. Shannon (1948), *A Mathematical Theory of Communication*, in «Bell System Technical Journal», n 27, pp. 379–423) described a textual analysis algorithm: «To construct [order-1 letter-level text] for example, one opens a book at random and selects a letter at random on the page. This letter is recorded. The book is then opened to another page and one reads until this letter is encountered. The succeeding letter is then recorded. Turning to another page this second letter is searched for and the succeeding letter recorded, etc. A similar process was used for [order-1 and order-2 letter-level text, and order-0 and order-1 word-level text]. It would be interesting if further approximations could be constructed, but the labor involved becomes enormous at the next stage».



Starting from Jurij Lotman's proposal of the 'semiosphere' (Lotman, 1990) as a conceptualization for cultural systems, based on the osmotic dynamic of the border and on the movement of peripheral elements from the side to the centre, and following Luigi Cavalli-Sforza's critic to a monolithic representation of languages, I have tried to configure a topological and relational space, animated by radiation and contraction movements.

## METHODOLOGY

World scale language data have been collected from UNESCO, Ethnologue, Scriptsource<sup>17</sup> and Jacques Leclerc's *L'aménagement linguistique dans le monde*<sup>18</sup>. They include different statistics about different aspects of languages and the use of languages (from the number of speakers to the writing system). These data have been processed with Gephi software to find out a representation of Global Language System, structured on a center-periphery dichotomy. The Fruchterman-Reingold Algorithm, where the sum of the force vectors determines which direction a node should move, while the step width, which is a constant, determines how far a node moves in a single step, has produced the first prototypical images.



The picture 1 represents the languages of the world connected with their writing system(s). The more central (with a galaxy of around 3,000 languages) is the *Latin script*, while thousands of not written languages form a sort of peripheral asteroid belt. The picture 2 represents the languages per country, revealing a polycentric political system of clusters where the periphery seems to be richer, in terms of variety, but less connected to the other elements of the system. The picture 3 shows the official languages connected with all the languages spoken in the country where they have an official status. It gives a possible representation of diglossic condition; in particular it reveals the interferences between the main polarities of English, Spanish and French.

The next step of the project aims at testing the validity of this relational-topological visualization at growing scales of analysis. It will be applied to the literary polisystem, to visualize its centers and peripheries (authors, books, literary languages), then to the works of single writers, to observe the core of his/her style, main topics, recurrent characters or places, discursive strategies. Finally, the aesthetic/ethic value of this center/periphery based iconic structure will be checked providing a discourse analysis of a small corpus of texts: the first case study has been individuated in the last anthology of Italian minority/dialectal writers.

---

<sup>17</sup> ScriptSource (available at <http://scriptsource.org>) is a dynamic, collaborative reference to the writing systems of the world, with detailed information on scripts, characters, languages - and the remaining needs for supporting them in the computing realm. It is sponsored, developed and maintained by SIL International.

<sup>18</sup> Leclerc, J., *L'aménagement linguistique dans le monde*, Québec, CEFAN, Université Laval, 25 novembre 2014, [<http://www.axl.cefan.ulaval.ca/europe/danemark.htm>].

## RESULTS

This is a currently underway research project: its main objective is to test and observe the relationship between the shape of the text and the theoretical framework of its investigation. The first achieved result, related to the visualization experiments of world scale language data, has been the possibility to organize a huge amount of information in an explorable centre based image, where some relevant nodes and connection have emerged. In this way, the concept of periphery and of border have been less theoretically and more empirically based.

However, there are a lot of open questions and hypothesis. The main expected result is to understand if the Lotman's proposal of topology as a metalanguage could effectively work through different scales and if and how it could affect the representation and perception of texts.

## BIBLIOGRAPHY

- [1] Andrews, B. (2003), *Electronic poetic*, in *The cybertext yearbook 2002*, Eskelinen and Koskimaa, editors, Jyväskylä.
- [2] Arnheim, R. (1974), *Il pensiero visivo*, Einaudi, Torino.
- [3] Arnheim, R. (1994), *Il potere del centro. Psicologia della composizione nelle arti visive*, Einaudi, Torino.
- [4] Borghi, A. M., Barsalou, W. L. (2002), *Prospettive nella concettualizzazione*, in Garotti, P., AIP Congresso nazionale della sezione di psicologia sperimentale, Nautilus, Bologna.
- [5] Cavalli-Sforza, L.L. (1996), *Geni, popoli e lingue*, Adelphi, Milano.
- [6] Cavalli-Sforza, L.L. e Feldman, M. (1981), *Cultural transmission and evolution. A quantitative approach*, Princeton UP, Princeton.
- [7] Cavalli-Sforza, L. L., Wang, W.S.Y. (1986), *Spatial distance and lexical replacement*, in "Language", LXII, 1.
- [8] Cramer, F. (2003), *Words made flesh. Code, Culture, Imagination*, Piet Zwart Institute, Rotterdam.
- [9] Di Napoli, G. (2011), *I principi della forma*, Einaudi, Torino.
- [10] Flusser, V. (2000), *Towards a Philosophy of Photography*, Reaktion books, London.
- [11] Gallese, V., Lakoff, G. (2005), *The brain's concepts: the role of the sensory-motor system in conceptual knowledge*, in "Cognitive neuropsychology", 21.
- [12] Gibson, J. J. (1999), *Un approccio ecologico alla percezione visiva*, Bologna, Il Mulino.
- [13] Lakoff, G. e Johnson, M. (2003), *Metaphors we live by*, Chicago University Press, London.
- [14] Leslie, J. (1995), *Ambiguity machines*, in «Ejournal», 5, 2.
- [15] Lotman, J. M. (2001), *Universe of the Mind: A Semiotic Theory of Culture (1990)*, trad. en. by A. Shukman, Tauris & C., London - New York.
- [16] Magritte, R. (1929), *Les mots et les images*, in «La Révolution surréaliste», V. 12.
- [17] McNeill, D. (1979), *The conceptual basis of language*, Routledge, New York.
- [18] Moretti, F. (2005), *Graph, maps, trees. Abstract Models For A Literary History*, Verso Books, London-New York.
- [19] Nelson, T. (1980), *Literary machines*, Mindful press, Sausalito.
- [20] Ong, W. (1982), *Orality and Literacy*, Routledge, London.
- [21] Thom, R., *Morfologia del semiotico* (1985), trad. it. di S. Costantini et al., Meltemi, Roma 2006.

# Un progetto per la consultazione on-line degli archivi del restauro (1850-1915).

Gabriella Guarisco<sup>1</sup>, Simonetta Ciranna<sup>2</sup>, Chiara Devoti<sup>5</sup>, Marco Felli<sup>2</sup>, Elena Fioretto<sup>1</sup>, Vittorio Foramitti<sup>3</sup>, Rossana Gabaglio<sup>1</sup>, Nora Lombardini<sup>1</sup>, Alessandra Maniaci<sup>3</sup>, Monica Naretto<sup>5</sup>, Giuseppina Perusini<sup>3</sup>, Elisa Piolatto<sup>5</sup>, Martina Visentin<sup>3</sup>  
[gabriella.guarisco@polimi.it](mailto:gabriella.guarisco@polimi.it), [simonetta.ciranna@univaq.it](mailto:simonetta.ciranna@univaq.it), [chiara.devoti@polito.it](mailto:chiara.devoti@polito.it),  
[marco.felli@graduate.univaq.it](mailto:marco.felli@graduate.univaq.it), [elena.fioretto@polimi.it](mailto:elena.fioretto@polimi.it), [vittorio.foramitti@uniud.it](mailto:vittorio.foramitti@uniud.it),  
[rossana.gabaglio@polimi.it](mailto:rossana.gabaglio@polimi.it), [nora.lombardini@polimi.it](mailto:nora.lombardini@polimi.it), [alessandra.maniaci@unirc.it](mailto:alessandra.maniaci@unirc.it),  
[monica.naretto@polito.it](mailto:monica.naretto@polito.it), [giuseppina.perusini@uniud.it](mailto:giuseppina.perusini@uniud.it), [elisa.piolatto@polito.it](mailto:elisa.piolatto@polito.it),  
[martina.visentin@uniud.it](mailto:martina.visentin@uniud.it)

1. Politecnico di Milano, 2. Università degli studi dell'Aquila, 3. Università degli studi di Udine, 4. Università degli Studi "Mediterranea" di Reggio Calabria, 5. Politecnico di Torino

## ABSTRACT

Il progetto di ricerca ha l'obiettivo di studiare, sistematizzare e pubblicare on-line le fonti archivistiche relative al restauro dei monumenti tra il 1850 e il 1915, tramite la loro digitalizzazione e inventariazione. L'importanza della ricerca riguarda la possibilità di divulgare e rendere accessibile la conoscenza della storia edilizia dei monumenti, spesso trascurata sia dal punto di vista della storia dell'arte e dell'architettura, sia da quello tecnico-professionale: l'accesso diretto ai documenti riguardanti i restauri permetterà infatti un significativo incremento della qualità della progettazione ed esecuzione dei lavori di conservazione dei monumenti e di intervento su tessuti storici della città.

Il metodo della ricerca prevede diverse fasi:

1. La ricognizione dei fondi archivistici delle istituzioni coinvolte negli interventi di restauro;
2. La scansione dei documenti conservati presso gli Archivi;
3. La predisposizione della piattaforma informatica;
4. L'indicizzazione dei documenti;
5. Il riversamento dell'indicizzazione e dei documenti nella piattaforma;
6. La pubblicazione on-line;
7. La pubblicazione di testi sui risultati dell'attività compiuta.

Completata la piattaforma informatica e provveduto alla pubblicazione on-line, il database potrà essere implementabile mediante l'inserimento di altre fonti documentarie.

Gli ambiti territoriali di interesse per il primo sviluppo sperimentale del progetto comprendono:

- il centro Italia a partire da Roma, con particolare attenzione alle trasformazioni delle architetture religiose nel corso del pontificato di Pio IX;
- il Friuli Venezia Giulia, con gli interventi seguiti dalle istituzioni italiane ed austriache preposte alla tutela dei monumenti;
- il Piemonte e la Valle d'Aosta con l'intensa attività dell'Ufficio Regionale per la Conservazione dei Monumenti di Piemonte, Liguria e Valle d'Aosta che vide come protagonista Alfredo d'Andrade;
- la Lombardia, con la vasta attività dell'Ufficio per la Conservazione dei Monumenti della Lombardia dove operarono Camillo Boito e Luca Beltrami.
- Ravenna, con i restauri dei principali monumenti seguiti da Corrado Ricci.
- La Sicilia.

## PAROLE CHIAVE

monumenti, restauro, documenti, digitalizzazione, consultazione.

## IL PROGETTO DI RICERCA

Il progetto di ricerca si prefigge l'obiettivo di studiare, sistematizzare, integrare e rendere accessibili allo studio il significativo patrimonio di fonti documentarie relative al restauro ed ai restauratori che, tra il 1850 e il 1915, effettuarono diffusamente la conoscenza e l'interpretazione di beni religiosi e civili, inaugurando una straordinaria stagione italiana per il restauro. Proprio quella fase avviò una riflessione profonda intorno al modo di intendere e praticare il restauro attraverso la quale si elaborarono teorie e metodologie codificate in Relazioni, Carte, contributi teorici che da lì in avanti vennero riconosciuti come fondanti a livello internazionale, e per i quali ancora oggi si guarda all'Italia come indiscusso luogo di riferimento disciplinare. La produzione storiografica italiana è vastissima in questo settore, ma le principali fonti documentarie sui restauri, ovvero i documenti scritti e iconografici che ne documentano scelte e approcci risultano per lo più indisponibili poiché solo parzialmente pubblicati e conservati in una moltitudine di fondi differenti presso Istituti come Archivi pubblici, compresi quelli di Stato, Gallerie e Musei, Soprintendenze, Biblioteche pubbliche, diffusi su tutto il territorio della penisola italiana

L'urgenza della sistematizzazione digitale delle fonti su descritte costituisce una priorità sia per la cultura sia per l'esecuzione di corretti cantieri di restauro sul patrimonio architettonico nazionale che, come recita il Codice dei Beni culturali del 2004 (la legge italiana per la tutela dei beni culturali), deve essere indirizzato alla conservazione del bene. La sua conservazione può avvenire solo nel rispetto della autenticità storico/materiale, che vuol dire riconoscibilità di tutti i periodi storici individuabili sul bene stesso.

L'importanza di questa operazione si allinea perfettamente con la Digital Agenda for Europe, parte delle strategie Europee del 2020, che intende utilizzare sfruttare i benefici derivanti dall'utilizzo degli ICT (*Information and Communication Technologies*) per la crescita economica dell'Europa, per lo sviluppo delle possibilità lavorative e per il miglioramento della qualità della vita dei cittadini. Tutto questo attraverso specifiche azioni informatizzate finalizzate a conservare e tramandare al futuro la memoria culturale, in tutte le sue forme, dell'Europa stessa.

Infatti, fin dall'agosto del 2006 la Commissione Europea preposta alla stesura delle raccomandazioni per la digitalizzazione, l'accessibilità on line dei documenti attestanti la cultura dell'Europa e la loro conservazione attraverso la digitalizzazione, ha invitato gli Stati membri a ottimizzare la potenzialità culturale ed economica del patrimonio culturale.

Il lavoro si articola attraverso la conoscenza e la catalogazione dei documenti d'archivio nelle diverse sedi in esame, con il fine di individuare, attraverso il confronto critico dei dati rilevati, precisi criteri di indicizzazione.

La comparazione fra i diversi archivi non ancora digitalizzati ed il network composto dai differenti conoscitori e studiosi di questi stessi archivi, ha la funzione ultima di determinare i parametri minimi necessari attraverso i quali identificare i documenti e descriverne i contenuti.

A questo proposito, la mancanza di un vocabolario convenzionale utile alla loro descrizione, così come, per esempio, *l'Art & Architecture Thesaurus – AAT* costruito dalla Getty Foundation con lo scopo di dare un significato univoco alla terminologia in uso nel settore dei Beni Culturali, o il ICOMOS-ISCS - *Illustrated glossary on stone deterioration patterns*, utile alla descrizione, ancora univocamente definita, dei degradi dei materiali lapidei, comporta la messa a punto di un nuovo vocabolario. Questo strumento, votato alla descrizione delle attività di restauro svolte su varie tipologie di monumenti (la chiesa, il resto archeologico, il

palazzo patrizio, e così via) nel periodo in oggetto, partendo dai glossari convenzionali, come appunto l'AAT, deve costituire uno strumento riconosciuto, accessibile e universalmente accessibile.

Questa operazione richiede non solo che si definiscano le *keywords* e la loro concatenazione (che costituiscono il metadato), ma che gli stessi descrittori (*keywords*) e gli specifici metadati siano opportunamente spiegati.

L'identificazione dei criteri è certamente prioritaria e, soprattutto, di non facile realizzazione. Essa parte, infatti, dalla coscienza che nell'ambito della cultura della conservazione esistono importantissime collezioni di dati (che riguardano, per esempio, la consistenza geometrica e materiale dei beni, e non, solamente, la storia dei loro restauri) e che i parametri adottati per la loro archiviazione li rendono praticamente inaccessibili, se non attraverso laboriose e a volte estenuanti ricerche.

Esiste, oggi, una importante discrepanza fra l'enorme quantità di testi a stampa del XVIII e XIX secolo (fonti notevoli e fondamentali nella ricerca della cultura del restauro nell'arco di tempo considerato) digitalizzati e disponibili su importanti e ben note piattaforme di divulgazione, come GOOGLE LIBRI e ARCHIVE.ORG, e la scarsa disponibilità di documenti di archivio digitalizzati e, quindi, immediatamente fruibili.

Questa diretta disponibilità renderà i dati storici utili sia nella ricerca scientifica, sia nella pratica del restauro e della conservazione del patrimonio architettonico, sia nella formazione degli addetti ai lavori (come liberi professionisti e funzionari della pubblica amministrazione) e dei ricercatori. Inoltre rientrerà nel delicato campo della valorizzazione del patrimonio culturale, perché (nel limite della "durata" del supporto fisico dei documenti digitalizzati e del formato di digitalizzazione) contribuisce alla conservazione del patrimonio stesso, e perché aiuterà la sua accessibilità e la sua divulgazione, sulla più ampia scala possibile (orizzontale, in senso geografico, e verticale, in termini di approfondimento dei vari gradi di conoscenza).

Non indifferente è il fatto che solo recentemente, dopo l'entrata in vigore della Legge 124/2017, è possibile la riproduzione con mezzi propri e la pubblicazione a fini scientifici dei beni culturali di qualsiasi tipo, ivi compresi i beni archivistici, senza il pagamento di alcun diritto o bollo, cosa che per anni ha reso più difficile la ricerca e la divulgazione, soprattutto in rete,

Molto spesso gli storici dell'architettura non si sono fatti carico delle trasformazioni prodotte nelle campagne di restauro, con il risultato di giungere talvolta ad attribuzioni clamorosamente sbagliate (per esempio l'attribuzione da parte di Robert Venturi della loggia di Palazzo Strozzi al Quattrocento, trattandosi invece di un intervento di restauro condotto da Giovanni Poggi nel 1938). Quello che interessava al folto gruppo di restauratori ottocenteschi era quel "valore storico" che reclamava a gran voce un ritorno all'indietro nel tempo, affinché quelle fabbriche fossero le testimonie fisiche di una storia e una cultura reperibile su tutto il territorio della nascente Italia unificata. Di contro, la storia "recente" delle trasformazioni delle fabbriche mediante i restauri ha prodotto un significativo mutamento nella consistenza fisica di questi monumenti che oggi, a più di cent'anni di distanza, risultano nuovamente degradati e dove talvolta si fatica a leggere direttamente sulle fonti materiali quelle modificazioni apportate più di cent'anni fa. Proprio in questo senso la ricerca intende portare in luce quanto avvenuto, quali furono i protagonisti, quale il dibattito, quali i risultati. Nel caso specifico, si tratta di mettere in evidenza la storia dei restauri condotti sul patrimonio architettonico quale parte integrante della storia della cultura nazionale e di rendere noto anche al grande pubblico che in Italia ci sono stati protagonisti al pari di Viollet-le-Duc in Francia, quali Alfredo D'Andrade, Camillo Boito, Giuseppe Fiorelli, Corrado Ricci ed altri che hanno svolto un lavoro oggi da conoscere e rivalutare. La digitalizzazione e divulgazione di questi documenti è uno dei processi della valorizzazione dei beni culturali.

L'elevato grado di integrabilità, la possibilità di collegamenti semantici fra porzioni testuali differenti, l'aggregazione delle informazioni e la diffusione dei contenuti hanno scopi di ricerca, divulgativi ed anche progettuali (si pensi, in questo caso, all'uso degli archivi storici nell'applicazione del BIM - *Building Information Modelling* agli edifici esistenti e al patrimonio culturale). Il riferimento al BIM non è del tutto casuale. Infatti rappresenta una delle più avanzate e, soprattutto cogenti, applicazioni degli ICT messa in atto dal mercato Europeo dell'edilizia nell'affidamento degli appalti pubblici che, di fatto, non riguardano solo il nuovo edificato ma, anche, quello esistente e, nello specifico, quello a carattere monumentale (il *Cultural Heritage* o, più precisamente, *Architectural Heritage*). [15]

Da tempo si sono applicati gli ICT alla gestione del patrimonio culturale. Già a partire della metà degli anni Novanta del secolo scorso, si è indagata l'applicazione del GIS (*Geographical Information System*) nella gestione sia dei centri storici (come sistema complesso di edifici - con differente struttura morfologia e strutturale- e connettivo urbano), sia dello stesso edificio, a prescindere dal suo valore "monumentale" (come sistema strutturale, costruttivo e storico, cioè costruzione che è risultato di manutenzioni, rifacimenti, adeguamenti e restauri succedutesi nel tempo).

Dal punto di vista tecnico-professionale l'accesso diretto ai documenti relativi ai restauri permetterà un significativo incremento e della qualità della progettazione ed esecuzione dei lavori di conservazione dei monumenti. Molto spesso, infatti, a causa della limitata conoscenza del patrimonio archivistico esistente - a volte non completamente inventariato - i progettisti dei lavori di restauro non sono nemmeno a conoscenza del fatto che il monumento sul quale intendono intervenire è già stato restaurato nel corso dell'Ottocento, e tanto meno sono informati degli interventi effettuati all'epoca anche dal punto di vista delle tecniche e dei materiali utilizzati.

Se si assume, come riferimento, l'applicazione delle Norme Tecniche per le Costruzioni (NTC) del 2008 (preposte al calcolo delle nuove strutture edilizie e della verifica delle esistenti) appare subito chiaro come il livello di conoscenza della fabbrica vada dichiarato dal progettista al momento dell'adozione del fattore di confidenza che influenza i parametri di risposta strutturale dell'edificio a fronte delle sollecitazioni esterne (come riportano le NTC al cap. 8, si tratta di «ulteriori coefficienti parziali di sicurezza che tengono conto delle carenze nella conoscenza dei parametri del modello»).

La possibilità di consultare i documenti permetterà quindi una più approfondita conoscenza ed anamnesi del monumento. In mancanza di questa conoscenza, può capitare che certe scelte progettuali ottocentesche, determinate dalla corretta applicazione pratica dei principi teorici del restauro allora seguiti, vengano completamente stravolte, cancellando la testimonianza materiale di un fondamentale momento nella cultura e nella vita dei monumenti poiché, come affermò De Angelis D'Ossat, «non si può considerare il restauro dei monumenti come un prodotto, anzi un sottoprodotto culturale tipico della civiltà moderna. Il fenomeno deve essere ricondotto nell'ambito continuo degli interventi architettonici su edifici esistenti che si sono realizzati in ogni età».

Oltre a ciò, il progetto prevede la creazione di un unico data base nel quale sarà possibile reperire le informazioni sui singoli manufatti ed i lavori di restauro che hanno subito, sulle istituzioni che hanno promosso e controllato i lavori, sui progettisti, sulle figure che hanno fornito pareri ed hanno contribuito al dibattito sul restauro, sugli operatori che hanno effettuato i restauri, sulle tecniche ed i materiali impiegati, con particolare attenzione per quelli derivanti dalle sperimentazioni e dai brevetti ottocenteschi ora non più utilizzati, come il metodo Pettekofer per la "rigenerazione" dei dipinti o il cemento idrofugo Ponti. La disponibilità di queste informazioni ed il loro collegamento incrociato fornirà un fondamentale aiuto alla ricerca ed alla didattica e consentirà il facile reperimenti di informazioni sul quadro generale delle istituzioni e sugli interventi di restauro eseguiti in

Italia e sulla storia degli edifici. Sarà inoltre possibile definire la effettiva attività dei protagonisti che, sia nel ruolo di progettisti, sia in quello di funzionari addetti al controllo dei lavori, sia per nel contributo al dibattito ebbero un ruolo fondativo nella disciplina del restauro in Italia. Oltre a questi, non indifferente risulta la ricognizione dell'attività dei restauratori nel territorio nazionale, facendo riferimento per esempio alle figure di Guglielmo Botti e di Antonio Bertolli nel campo del restauro degli affreschi.

Gli ambiti geografici di interesse per il primo sviluppo sperimentale del progetto comprendono diverse aree significative del territorio nazionale. La significatività dipende anche dalla varietà di monumenti presenti nei singoli comprensori proposti e la diversità di attori in azione.

Nel periodo considerato, 1850-1915, è costruita la scienza del restauro e la legislazione di tutela. Queste informazioni devono essere sintetizzate, tradotte in termini univocamente definiti e indicizzate. La grande varietà sia di tipologie di monumenti sia di criteri di intervento utilizzati offre una occasione interessante e strumentale all'individuazione di parametri generalizzati e, per questo, più facilmente condivisibili.

L'Italia è il territorio analizzato e utilizzato come area storico-geografica campione, nella coscienza che l'analisi dell'ampia varietà dei monumenti italiani e delle tecniche e metodologie di restauro adottate possano costituire un ottimo punto di partenza e di paragone. Le regioni che saranno prese in considerazione nella prima fase del progetto saranno:

- il centro Italia a partire da Roma, con particolare attenzione all'opera dell'architetto Virginio Vespignani, alle trasformazioni delle architetture religiose a partire da quelle realizzate nel corso del pontificato di Pio IX;
- il Friuli Venezia Giulia, di particolare interesse per la compresenza, fino alla fine della Grande Guerra, delle istituzioni italiane ed austriache preposte alla tutela dei monumenti;
- il Piemonte e la Valle d'Aosta, a partire dal *milieu* culturale che originò la scoperta del romanico e in particolare la figura di Edoardo Arborio Mella, fino all'intensa attività dell'Ufficio Regionale per la Conservazione dei Monumenti di Piemonte, Liguria e Valle d'Aosta, che vide come protagonista Alfredo d'Andrade;
- la Lombardia, con la grande attività svolta principalmente sul romanico dall'Ufficio per la Conservazione dei Monumenti della Lombardia e dove operarono Camillo Boito e Luca Beltrami;
- Ravenna, con i restauri dei principali monumenti tardo-romani e bizantini seguiti da Corrado Ricci, fautore della prima Soprintendenza ai Monumenti italiana;
- La Sicilia, anche con il suo straordinario patrimonio archeologico.

Per quanto concerne lo stato dell'arte, su un fronte il riferimento è rappresentato dall'amplissima storiografia sul restauro, in particolare quella che indaga la storia e critica dei restauri. Nel campo della digitalizzazione e messa in accessibilità di archivi si citano la piattaforma francese *Gallica* della BNF, la base dati *Mérimée* sui Monumenti Nazionali francesi ([www.mediathèque-patrimoine.culture.gouv.fr/pages/recherche/bases/merimee.html](http://www.mediathèque-patrimoine.culture.gouv.fr/pages/recherche/bases/merimee.html)), i portali italiani SAN e SIUSA ([san.beniculturali.it/web/san/home](http://san.beniculturali.it/web/san/home); [siusa.archivi.beniculturali.it](http://siusa.archivi.beniculturali.it)) e lo strumento ARISTOS parte del più ampio progetto ARTPAST del MIBACT (<http://www.artpast.iccd.beniculturali.it/>). Attuali tentativi di pubblicazione on line di patrimoni tematici sono rappresentati dal progetto di catalogo accessibile dei beni dell'eredità Rothschild dell'INHA e del Louvre ([collections.rothschild.inha.fr](http://collections.rothschild.inha.fr)). Da citare anche l'attività di riproduzione e catalogazione dei documenti già avviata dal Laboratorio Informatico per la Documentazione Storico Artistica (LIDA) dell'Università degli studi di Udine.

Il metodo della ricerca prevede diverse fasi operative:

1. La ricognizione dei fondi archivistici delle istituzioni coinvolte negli interventi di restauro;
2. La scansione dei documenti cartacei conservati presso gli Archivi pubblici e/o privati;
3. La predisposizione della piattaforma informatica;
4. L'indicizzazione dei documenti;
5. Il riversamento dell'indicizzazione e dei documenti nella piattaforma;
6. La pubblicazione on-line;
7. La pubblicazione di testi sui risultati dell'attività compiuta e sui contributi alla conoscenza dei manufatti e dei protagonisti.

In particolare, ai fini degli obiettivi della ricerca, i metadati da utilizzare per la descrizione dei documenti e per la loro indicizzazione dovranno comprendere principalmente: regione, provincia e comune del manufatto; archivio di provenienza dei documenti; nome del manufatto; data; istituzioni coinvolte nei lavori; nomi degli autori del progetto e dei pareri e degli operatori del restauro; se rilevanti tecniche o materiali utilizzati nel restauro.

Una volta completata la piattaforma informatica, verificata l'indicizzazione e pubblicate, on-line, le istruzioni per la consultazione, la piattaforma stessa potrà essere implementabile mediante l'inserimento di archivi o parti di essi già indicizzati e scansionati da Enti diversi. Il fine è quello di realizzare una piattaforma facilmente manutenibile, implementabile e compatibile. Questo proprio per venire incontro all'esigenza di non disperdere i dati che si vanno ad implementare e quelli che già sono in nostro possesso, assicurandone un continuo aggiornamento e facilitandone l'utilizzo.

La consultazione on-line degli Archivi del restauro potrà quindi divenire uno strumento straordinario per la ricerca relativa alla storia dei singoli monumenti e siti ed alle opere degli architetti e restauratori che operarono nella neonata nazione italiana, a partire dal momento del concepimento della sua unificazione.

## BIBLIOGRAFIA

Testi sulla storia del restauro nelle aree considerate

- [1] Cerri, M.G., Biancolini Fea, D., Pittarello, L. 1981. *Alfredo d'Andrade. Tutela e restauro*, Firenze, Vallecchi.
- [48] Ciranna, S. 2007. *I Martinori. Scalpellini, inventori, imprenditori dalla città dei papi a Roma capitale*, Roma, Camera di Commercio.
- [49] Ciranna, S. Doti, G. Neri, M. L. 2011. *Architettura e città nell'Ottocento. Percorsi e protagonisti di una storia europea*, Roma, Carocci.
- [50] Foramitti, V. 2004. *Tutela e restauro dei monumenti in Friuli - Venezia Giulia, 1850-1915*, Udine, Edizioni del Confine.
- [51] Giusti, M. A., Naretto M. (a cura di) 2014. *Arte di Conservare / Conservare con l'Arte*. Castello, villa, villeggiature d'artisti a Rivara. Pisa, ETS.
- [52] Guarisco, G. (a cura di) 2015. *Fernand de Dartein e l'architettura Romanica comasca. Viaggio in un archivio inesplorato*. Ariccia (RM):Ermes.
- [53] Guarisco, G. (a cura di) 2012. *Fernand de Dartein. La figura, l'opera, l'eredità*, Firenze, ALINEA.
- [54] Mozzo, M. Visentin, M. 2014. *Vicende di tutela in Friuli tra Ottocento e Novecento. Monumenti, oggetti d'arte e paesaggio*, Udine, Forum.
- [55] Oteri, A. M. 2002. *Riparo, Conservazione restauro nella Sicilia orientale, o del "diffinitivo assetto" 1860-1902*, Roma, Gangemi.
- [56] Perusini, G. Fabiani, R. (a cura di) 2014. *La conservazione dei monumenti e delle opere d'arte in Friuli nell'Ottocento*, Udine, Forum.



- [57] Perusini, G. Fabiani, R. (a cura di) 2008. Conservazione e tutela dei beni culturali in una terra di frontiera: il Friuli Venezia Giulia fra Regno d'Italia e Impero Asburgico (1850 – 1918), Vicenza, Terra Ferma 2008.
- [58] Picone, R. 2012. Il restauro e la questione dello stile, il secondo Ottocento nel mezzogiorno d'Italia, Napoli, Editrice Politecnica.
- [59] Prescia, R. 2012. Restauri a Palermo, architettura e città come stratificazione, Palermo, Kalòs
- [60] Testi su ICT e beni culturali
- [61] Achille, C. Lombardini, N. Valentini, M. 2013. *Sistemi di protezione per le aree archeologiche: linee guida in tempo di crisi*, in *Conoscere, Conservare, Valorizzare*, a cura di A. R. Genovese, Napoli, Arte tipografica Editrice, pp. 393-417.
- [62] Achille, C. Lombardini, N. Tommasi, C. 2015. BIM and cultural heritage: compatibility tests in an archaeological site, in *Building Information Modelling (BIM) in Design, Construction and Operations*, Editors L. Mahdjoubi, C.A. Brebbia, R. Laing, Ashurst Southampton, WIT Press, pp. 593-604.
- [63] *Predisposizione di un sistema informativo territoriale finalizzato al progetto di conservazione dei centri storici*, responsabile scientifico nazionale prof. Marco Dezzi Bardeschi, finanziamento ricerca nazionale fondi MURST, periodo 1998-2000.
- [64] Gasperi, D. 2013. *Il patrimonio partecipato. L'ICT per i Beni culturali* in Galofaro C., Gasperi D., Proni G., Ragonese R., *Semiotica e ICT per i Beni Culturali*, Milano, Franco Angeli, pp. 89-116
- [65] Paolini, P. Di Blas, N. Alonzo, F. 2005. *ICT per i beni culturali, esempi di applicazione*, in *Mondo Digitale*, n. 3, pp. 44-61  
[http://people.unica.it/psanjust/files/2008/01/ict\\_per\\_i\\_beni\\_culturali-paolini.pdf](http://people.unica.it/psanjust/files/2008/01/ict_per_i_beni_culturali-paolini.pdf)  
 (consultato novembre 2018)
- [66] Doerr, M. Gradmann, S. Hennicke, S. Isaac, A. Meghini, C. van de Sompel, H. 2010. *The European data model (EDM)*, in *76th IFLA General Conference and Assembly*, pp. 1-12  
<https://www.ifla.org/past-wlic/2010/149-doerr-en.pdf> (consultato novembre 2018)
- [67] Commission recommendation of 27 October 2011 on the digitisation and online accessibility of cultural material and digital preservation (2011/711/EU), Official Journal of the European Union, 20.10.2011, L. 283/39 <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:283:0039:0045:EN:PDF>  
 (consultato novembre 2018)

# **Strumenti di Laboratorio. Il Laboratorio Informatico per la Documentazione Storico Artistica del Dipartimento di Studi Umanistici e del Patrimonio Culturale dell'Università di Udine tra ricerca e didattica.**

**Martina Visentin**

Università degli Studi di Udine

## **ABSTRACT**

L'intervento intende illustrare le attività del Laboratorio LIDA del Dipartimento di Studi Umanistici e del Patrimonio Culturale dell'Università di Udine in particolare riferimento alle applicazioni della piattaforma di gestione dati E-Dvara per la ricerca e la didattica.

Diretto da Donata Levi, il LIDA (Laboratorio Informatico per la Documentazione Storico Artistica) è attivo dal 2006 e nasce quindi in anni piuttosto precoci. Con questa proposta di intervento, Martina Visentin (impegnata nel LIDA come tecnico di laboratorio laureato e già assegnista di ricerca) propone di illustrare le esperienze maturate (da umanista) nello sviluppo di banche dati dedicate a progetti di ricerca anche di altre istituzioni (Memofonte di Firenze, Scuola Normale Superiore di Pisa, Università di Lecce, Università di Urbino, Università di Firenze, Soprintendenza di Firenze...).

Strumento di tali collaborazioni è stato in particolare il SW E-Dvara sviluppato dal dipartimento di Matematica dell'Ateneo per la messa a punto di uno strumento web based di gestione e restituzione dati in grado di rispondere alle esigenze di fruizione dei dati relativi alla ricerca in ambito storico artistico e storico critico (e-contents platform for Heritage and Science <http://edvara2.uniud.it/>). Negli anni il Laboratorio ha applicato la piattaforma E-Dvara nell'ambito di attività di ricerca di diversi docenti del Dipartimento per la strutturazione di sistemi relazionali di dati testuali e visivi relativi alla storia dell'arte, della critica e del collezionismo. Parallelamente, il SW è stato usato anche con funzione didattica e di coinvolgimento degli studenti che attraverso tale esperienza hanno avuto così la possibilità di prendere parte – a seconda dei progetti - alla progettazione di una struttura di relazione dati come alle fasi di popolamento.

Con il suo intervento Martina Visentin intende illustrare tali esperienze del Laboratorio, anche rendendo conto della progettazione in corso di un nuovo strumento di gestione dati in grado di mantenere le stesse funzioni di E-dvara ma con modalità più aggiornate sia per le applicazioni nella ricerca che nella didattica.

## **PAROLE CHIAVE**

Cultural Heritage, Art History, history of art criticism, E-platoform, Data base

## **INTRODUZIONE**

Il software E-Dvara, sviluppato in collaborazione con il Dipartimento di Matematica dell'Ateneo, è stato applicato nell'ambito di attività di ricerca di diversi docenti del Dipartimento di Studi Umanistici e del Patrimonio Culturale dell'Università di Udine per la strutturazione di sistemi relazionali di dati testuali e visivi relativi alla storia dell'arte, della critica e del collezionismo. Parallelamente, dopo un prima fase di test, il SW è stato usato anche con funzione didattica e con il coinvolgimento degli studenti, in particolare nell'ambito

dei corsi di Storia della storia dell'arte, offrendo così agli studenti l'esperienza non comune in un percorso umanistico di comprendere le diverse possibilità di creare relazioni tra i dati.

## **PAGINA, TESTO E BIBLIOGRAFIA**

Il laboratorio LIDA (Laboratorio Informatico per la Documentazione Storico Artistica), diretto da Donata Levi, è attivo dal 2006 nel Dipartimento di Studi Umanistici e del Patrimonio Culturale dell'Università di Udine. Erede in particolare delle prime sperimentazioni di Paola Barocchi a Pisa e Firenze di applicazione degli strumenti informatici al trattamento dei dati relativi alla ricerca storico critica, il Laboratorio nasce quindi in anni piuttosto precoci. Nella penisola tarda infatti lo sviluppo di una riflessione sulle possibilità di gestione dati offerta alla ricerca storico artistica e storico critica da parte delle così dette Digital Humanities. In maniera piuttosto singolare inoltre il Laboratorio si è dedicato nel tempo sia al trattamento delle fonti per la ricerca, sia a progetti finalizzati alla divulgazione di contenuti scientifici (in particolare <http://infobc.uniud.it/>), mostrando così un'attenzione non scontata per le iniziative di comunicazione della ricerca.

Insieme a tale contesto di sviluppo, da subito il Laboratorio LIDA sperimenta una specifica competenza nello sviluppo di strumenti web per la gestione dei dati, come dimostra anche il ruolo di progettazione svolto nell'ambito del progetto ARTPST del Ministero per i Beni e le Attività Culturali e del Turismo, con lo sviluppo di ARISTOS (data base relazione finalizzato alla descrizione dei materiali documentari degli archivi storici e correnti delle Soprintendenze) e SICAR (sistema di georeferenziazione dei dati di restauro in corso): da una parte esperienze di stretta collaborazione con esperti informatici e dall'altra esperienze di gestione di materiali documentari tra loro molto diversi anche per l'appartenenza a periodi storici vari e articolati.

Nell'ambito del Convegno annuale AIUCD 2019, Martina Visentin (storico dell'arte con PhD in storia della critica, ora tecnico di laboratorio laureato impegnata presso il LIDA e a lungo assegnista di ricerca) propone di illustrare le esperienze del Laboratorio nello sviluppo di banche dati dedicate a progetti di ricerca condotti anche in collaborazione con altre istituzioni (Memofonte di Firenze, Scuola Normale Superiore di Pisa, Università di Lecce, Università di Urbino, Università di Firenze, Soprintendenza di Firenze...). Si tratta di strutture di dati sviluppate grazie in particolare alla piattaforma di gestione dati E-Dvara, strumento sviluppato dall'Università di Udine per la gestione di dati relativi alla ricerca umanistica.

Già a partire dell'avvio delle sue attività nel 2006, il Laboratorio ha infatti collaborato con il dipartimento di Matematica dell'Ateneo, per la messa a punto di uno strumento web based di gestione e restituzione dati in grado di rispondere alle esigenze di gestione dei dati relativi alla ricerca in ambito storico artistico e storico critico (e-contents platform for Heritage and Science <http://edvara2.uniud.it/>). Esito di tale collaborazione è stato il SW E-Dvara che permette un'autonoma strutturazione di base dati relazionali anche da parte del non esperto informatico con la gestione di tre ambienti di lavoro: archivi, templates, vocabolari. In particolare nella sezione template, l'utente può sviluppare singoli sistemi di relazione dati all'interno dei quali predisporre campi a testo libero, campi relazione, vocabolario, link, campi per il caricamento di immagini o di file di testo in diverso tipo di formato.

Grazie a tali strumenti di modellazione (del tipo di dati ritenuti significativi, come delle relazioni che intercorrono tra essi), il SW E-Dvara ha permesso al Laboratorio di sviluppare senza costi tecnici aggiuntivi data base utili sia in fase di sviluppo di ricerca che in fase di pubblicazione e divulgazione. Per altro il SW – fornito di un sistema di back end e di una

interfaccia per la modellazione del front end – permette di mostrare solo parte dei dati inseriti, consentendo così la possibilità di non divulgare informazioni (ad esempio) work in progress. Attraverso alcuni significativi esempi si propone una riflessione sull’esperienza maturata dal Laboratorio, a partire dalla singolare opportunità di affiancare i ricercatori informatici nello sviluppo di una piattaforma flessibile nei suoi usi e destinata in maniera specifica al patrimonio storico, artistico e culturale, sia nelle funzioni della ricerca che in quelle della didattica.

Anche grazie alle suggestioni legate alla partecipazione ai già citati progetti nazionali (in particolare ARTPST del MIBACT con lo sviluppo di ARISTOS e SICAR), il Laboratorio ha infatti sviluppato anche una serie di iniziative finalizzate al coinvolgimento degli studenti. La piattaforma E-Dvara è stata quindi applicata come occasione di formazione per gli studenti nel senso di un primo contatto con la strutturazione di un sistema di relazione tra dati, di loro organizzazione per lo studio o per attività di divulgazione.

Il SW E-dvara prevede infatti una sezione front-end che consente di modellare anche la visualizzazione dei dati inseriti attraverso un sistema di consultazione verso l’utente esterno. A titolo di esempio si vedano tra gli altri i seguenti progetti: <http://edvara2.uniud.it/marchetti/> (per la valorizzazione di un fondo fotografico in collaborazione con i Civici Musei di Udine con la creazione di due diversi itinerari, progetto sviluppato nell’ambito di un assegno di ricerca finanziato); <http://edvara2.uniud.it/accademiaverona/> (per la descrizione dell’archivio storico dell’Accademia di Belle Arti di Verona nell’ambito di una tesi di Laurea); <http://edvara2.uniud.it/municipalia-forli/> (per una ricerca sulle fonti edite e inedite per la ricostruzione della percezione del patrimonio e la sua tutela), <http://edvara2.uniud.it/esposizioni-arte-antica/> (per una ricerca svolta dall’Università degli Studi di Urbino sulle mostre d’arte antica organizzate tra l’Unità e la seconda guerra mondiale); <http://edvara2.uniud.it/ciol-fotografia-pellegrino/> (per la valorizzazione di una sezione della fototeca della Dipartimento di Studi Umanistici del Patrimonio Culturale del quale il Lida ha la gestione); <http://edvara2.uniud.it/tutelaFVG/> (per la descrizione di una fondo storico inedito con un’importante campagna di ricerca che ha condotto alla pubblicazione M. Mozzo, M. Visentin, *Vicende di tutela in Friuli tra Ottocento e Novecento. Monumenti, oggetti d’arte e paesaggio*, Udine, Forum, 2014).

Con il suo intervento Martina Visentin intende illustrare tali esperienze del Laboratorio, anche rendendo conto della progettazione in corso di un nuovo strumento di gestione dati in grado di mantenere le stesse funzioni di E-dvara ma con modalità più aggiornate sia per le applicazioni nella ricerca che nella didattica.

## BIBLIOGRAFIA

- [1] M. Visentin, Il LIDA dell’Università di Udine: esperienze di applicazione ICT per i beni culturali tra ricerca e valorizzazione. In: (a cura di): Auriemma R, *La democrazia della conoscenza. Patrimoni culturali, sistemi informativi e open data: accesso libero ai beni comuni?*, atti del convegno di Trieste (28-29 gennaio 2016), Forum Editrice Universitaria Udinese, Udine, 2017, pp. 187-197.
- [2] M. Visentin, Il LIDA\_FOTOTECA Esperienze per le Digital Humanities, dalla ricerca alla comunicazione, in *Le risorse digitali per la storia dell’arte moderna in Italia. Progetti, ricerca scientifica e territorio*, a cura di F. Conte, Roma, Edizioni di Storia e letteratura, 2018, pp. 39-54.

# Fondazione Memofonte: l'informatica «aiuta a razionalizzare la ricerca»

Martina Nastasi<sup>1</sup>, Donata Levi<sup>2</sup>

<sup>1</sup>Fondazione Memofonte onlus, Italy; [info@memofonte.it](mailto:info@memofonte.it)

<sup>2</sup>Fondazione Memofonte onlus, Italy

Il cammino della Fondazione Memofonte è iniziato nel 2000 come associazione creata da Paola Barocchi –Fondazione dal dicembre del 2006 –, ma affonda le sue radici in un passato più lontano e solido, costituito dall'attività pionieristica della fondatrice presso la Scuola Normale Superiore di Pisa nell'ambito dell'informatica applicata ai beni culturali. Negli anni in cui l'informatica iniziava a varcare i confini italiani con tutto il bagaglio di potenziali innovazioni, Paola Barocchi ne colse a pieno l'importanza in relazione alla ricerca umanistica, riuscendo a vederne gli sviluppi futuri e la nuova linfa vitale che tali strumenti erano in grado di apportare alla conoscenza. Come ebbe modo di dire in un'intervista del 1994, l'informatica «aiuta a razionalizzare la ricerca»: come tutti gli strumenti occorre però fare una buona taratura, pensare una «struttura molto chiara», sapere cosa si vuole, avere chiari degli obiettivi per evitare di andare fuori strada.

La Fondazione Memofonte si occupa prevalentemente dell'elaborazione informatica delle fonti storico-artistiche, vale a dire il trattamento digitale dei testi, siano essi manoscritti o a stampa, che afferiscono alla storia dell'arte. Le differenti tipologie di documenti vengono analizzate, trattate e strutturate nel rispetto delle caratteristiche precipue, cercando di offrire il sistema più razionale e funzionale per gli utenti. Primo grande progetto è stato il lavoro di trascrizione informatica delle *Vite di Pittori* di Francesco Maria Niccolò Gabburri, eseguito su commissione della Biblioteca Nazionale al fine di salvaguardare il manoscritto originale devastato dall'alluvione del 1966. Uno scopo conservativo supportato dalla possibilità di eseguire una sorta di restauro digitale: Paola Barocchi era infatti in possesso di un microfilm completo dei volumi allo stato prealluvione, perciò integro in tutte le sue parti. Numerosi altri lavori hanno fatto del sito della Fondazione Memofonte un punto di riferimento per la comunità degli studiosi, dimostrando quanto già negli anni Ottanta la sua fondatrice aveva intuito. Tutti i progetti di ricerca sono basati su solidi percorsi scientifici, che hanno permesso di strutturare database e/o trascrivere documenti manoscritti o a stampa, in modo da poter guidare gli utenti attraverso itinerari atti a facilitare la consultazione e la ricerca.

Si potrebbe schematizzare il contenuto del sito seguendo la suddivisione tematica della pagina dedicata alla ricerca:

- Arte e lingua: la sezione presenta i progetti frutto della collaborazione con l'Accademia della Crusca. I *corpora* testuali sono stati digitalizzati e sono consultabili secondo più modalità, permettendo interrogazioni anche tramite lemmari.
- Bibliografie e biblioteche d'arte: la consultazione delle prime bibliografie storiche presenti nei grandi manuali dei secoli XVII-XX, consente di reperire le fonti più accreditate del passato fruibili da eminenti studiosi e quindi molto significative per la ricostruzione storica relativa ad artisti, nuclei geografici, collezioni e istituzioni. La volontà di creare questa sezione nasce da ldesiderio di integrare la bibliografia dello Schlosser la quale oggi presenta i limiti di una preoccupazione qualitativa che ha escluso testimonianze di grande significato culturale.
- Collezionismo: inventari, cataloghi, descrizioni e carteggi offrono materiali per lo studio di nuclei e di tradizioni collezionistiche, come quella medicea, lorenese ed estense. La raccolta e la pubblicazione dei documenti della sezione è strettamente

legata agli studi barocchiani, che nel tempo si sono indissolubilmente legati al trattamento informatico.

- Guide, descrizioni e appunti di viaggio: la sezione, dedicata a quel genere letterario che in epoca moderna vede i suoi prodromi in ambito fiorentino e romano, con la pubblicazione delle guide di Roma e di Firenze ad opera di Francesco Albertini, presenta una cospicua raccolta di testi di guide e di descrizioni di complessi monumentali o museali. Organizzata topograficamente, tale raccolta per la ricchezza delle testimonianze (in special modo quelle napoletane e fiorentine) permette di seguire lo sviluppo di questa modalità di ricognizione sul territorio.
- Scritti d'arte XVI-XVIII secolo: Da Michelangelo Buonarroti a Giorgio Vasari, da Filippo Baldinucci a Luigi Lanzi la sezione ospita digitalizzazioni di testi fondamentali della storiografia artistica e dell'antiquaria (glittica e medaglistica), trascrizioni di documenti inediti e banche dati relativi a carteggi e cataloghi.
- Scritti d'arte XIX-XX secolo: testimonianze di varia natura (carteggi, trattati, discorsi, memorie, pubblicazioni) scritte da pittori, scultori e altri importanti protagonisti delle arti figurative, che grande fortuna hanno avuto tra l'Ottocento e il Novecento. La trattazione informatica consente un'agile valorizzazione degli infiniti riferimenti a persone, opere, luoghi, eventi, permettendo di tracciare un tessuto informativo utile per la ricostruzione delle complesse dinamiche storico culturali. Ogni personaggio ha una propria sottosezione che ne raccoglie i documenti a esso pertinenti, accessibili sia con ricerca monografica, che con verifiche complessive tramite un database generale.

I risultati delle ricerche trovano spazio anche nella rivista della Fondazione, «Studi di Memofonte», che dal 2008 accoglie studi nel campo dell'informatizzazione del patrimonio culturale e nelle ricerche storico-artistiche particolarmente incentrate sull'utilizzo delle fonti.

Il prodotto digitale ha qualcosa di simile a una pianta che ha bisogno di essere curata nel tempo perché sopravviva. E in questo caso si tratta di un tempo che sembra aver preso un ritmo vertiginoso, in cui la velocità di memoria futurista ha già iniziato a lasciare dietro di sé carcasse di creature diventate troppo presto obsolete e prive di nuova linfa in grado di farle sopravvivere. Gli strumenti informatici consultabili online sono frutto di duro lavoro e di investimenti economici, ma il carattere intermittente del flusso dei finanziamenti (spesso pubblici!) fa sì che dopo un dispendio generale di soldi ed energie non ci sia più la possibilità di curare siti, database, programmi, che perciò non vengono aggiornati.

L'impegno della Fondazione è dunque quello di preservare il suo patrimonio materiale (biblioteca, archivio e sede) ma anche di garantire un aggiornamento continuo di quanto presente sul sito., recentemente aggiornato sempre seguendo un aspetto significativo dell'attività di Paola Barocchi, vale a dire la volontà di semplificare l'apporto informatico, di usarlo nella sua elementarità. Si potrebbe quasi parlare di adozione di strumenti 'domestici', di pari passo con il rifiuto di cedere alle lusinghe di effetti speciali o di sistemi troppo complicati, quasi a esaltare il valore comunicativo della pubblicazione online.

# The role of research infrastructures in the area of DH education, training and competence building

Tanja Wissik

Austrian Academy of Sciences, AT, tanja.wissik@oeaw.ac.at

## ABSTRACT

The growing interest in digital humanities research has resulted in an increasing number of training events, individual courses, and degrees covering a broad range of topics at the cross-section of the humanities and ICT-based methods. The key players in the teaching and training landscape in the area of the digital humanities are universities and other higher education institutions. However, European research infrastructures such as DARIAH-EU (Digital Research Infrastructure for the Arts and Humanities) and CLARIN-ERIC (Common Language Resources and Technology Infrastructure) as well as regional DH associations are also contributing to the area of DH education, training and competence building. In this contribution, we will describe the role of research infrastructures for the development of DH training and education and showcase two initiatives that have been developed and promoted in the context of these research infrastructures, namely the Digital Humanities Course Registry and the #dariahTeach platform for open educational resources.

## KEYWORDS

Research infrastructures, education, training, competence building, DH teaching and training landscape

## INTRODUCTION

The growing interest in digital humanities research has resulted in an increasing number of training events, individual courses and degrees covering a broad range of topics at the cross-section of the humanities and ICT-based methods. The key players in the teaching and training landscape in the area of the digital humanities are universities and other higher education institutions. Across the world, they are developing DH curricula and delivering DH programmes, usually based on local strategies and available resources. However, European research infrastructures such as DARIAH-EU<sup>19</sup> (Digital Research Infrastructure for the Arts and Humanities) and CLARIN-ERIC<sup>20</sup> (Common Language Resources and Technology Infrastructure) as well as regional DH associations are also investing in DH training and education. In this contribution, we will describe the role of research infrastructures for the development of DH training and education and showcase two initiatives that have been developed and promoted in the context of these research infrastructures, namely the Digital Humanities Course Registry and the dariahTeach platform for open educational resources.

## ROLE AND USE OF RESEARCH INFRASTRUCTURES FOR EDUCATION, TRAINING AND COMPETENCE BUILDING

What role can these research infrastructures play in the transfer of knowledge and knowhow regarding digital methods and in the education and training of a new generation of young researchers? Most definitions of research infrastructures focus primarily on the technical

---

<sup>19</sup> [www.dariah.eu](http://www.dariah.eu)

<sup>20</sup> [www.clarin.eu](http://www.clarin.eu)

aspects of infrastructures. These definitions do not acknowledge that research infrastructures also function as central nodes of communication and expertise, which carry great potential for the education and training of young DH researchers, especially outside typical classroom settings. Other definitions, such as the one listed in the Austrian Infrastructure Action Plan, underline the importance of research infrastructures in education and training: “Forschungsinfrastruktur ist aber nicht nur technische Ausstattung für die Forschung, sondern auch ein Knotenpunkt für Kommunikation und essentiell für die Ausbildung von wissenschaftlichem und technischem Nachwuchs<sup>21</sup>“ (Österreichischer Forschungsinfrastruktur-Aktionsplan 2014-2020, 2014).

Furthermore, in the last ESFRI Roadmap Strategy Report on Research Infrastructures it was described, that “RIs also have a broad impact on skills and education agendas. Increasing the competences of staff, researchers and academic users, and through outreach to pupils and students, as well as the general public, they steadily improve the perception and understanding of science and technology in society” [2].

## DIGITAL HUMANITIES COURSE REGISTRY

Increasing the visibility of DH training activities is a major concern of the DH community and the Digital Humanities Course Registry was conceived to address this. It is designed as a hub to collect information on digital humanities programmes, courses, lectures and summer schools (see Figure 1).

Status	Course Name	Education Type	Institution	Department	Information	Curriculum
Not reviewed 2018-11-09	Double licence Lettres – Informatique	Degree: Bachelor Programme	Université Paris-Sorbonne	UFR Littérature française et comparée	<a href="#">Info</a>	<a href="#">Guide</a>
Not reviewed 2018-11-09	Heritage Visualisation	Degree: Master Programme	The Glasgow School of Art	School of Simulation and Visualisation	<a href="#">Info</a>	-
Not reviewed 2018-11-09	Master Humanités numériques	Degree: Master Programme	Université Lyon 2	ICOM	<a href="#">Info</a>	-
Not reviewed 2018-11-09	Introduction aux Humanités Numériques	Credits: Course	Le Mans Université	Département d'allemand	<a href="#">Info</a>	-
Not reviewed 2018-11-09	Pratiques numériques dans l'enseignement en France et en Allemagne	Credits: Course	Le Mans Université	Département d'allemand	<a href="#">Info</a>	-
Not reviewed 2018-11-09	Urheberrecht im europäischen digitalen Raum	Credits: Course	Le Mans Université	Département d'allemand	<a href="#">Info</a>	-
Not reviewed 2018-11-09	Comprendre et utiliser les bases de données	Credits: Course	Le Mans Université	Département d'allemand	<a href="#">Info</a>	-

Figure 1: Search interface of the DH Course Registry

The database contains, at this point, over 200 courses and training events in Europe and beyond (see Table 1). The search functionality is enhanced by the addition of an interactive map, making the listings searchable by location. The DH Course Registry is a joint effort of DARIAH-EU and CLARIN ERIC.

<sup>21</sup> However, research infrastructures are not only technical equipment for research, but also a hub for communication and essential for the education of young scientists and engineers (translation by the author).



Country	Number of Courses
Germany	33
Netherlands	29
France	25
Switzerland	23
Italy	16
Austria	12
United Kindom	11
Russian Federation	9
United States of America	8
Finland	7
Belgium	5
Spain	5
Ireland	4
Greece	4
Norway	3
Hungary	3
Israel	3
Croatia	2
Latvia	1
Czech Republic	1
Denmark	1
Sweden	1
Lithuania	1
Poland	1
Slovenia	1
Portugal	1

Table 1: Courses per country

In the year 2018 so far (up until 10 Nov), the DH Course Registry website received 3597 visits from 90 different countries. Most of the visits (2873) have been from countries in Europe. The DH Course Registry targets various user groups: students, lecturers, researchers, programme administrators and decision makers.

Usually, such collections of information or databases are at high risk to become outdated very fast. In order to prevent this and to monitor the quality of the registry, the project relies on national moderators for each country who are tasked with curating the content entered by individual lecturers or program directors [1]. In order to remain sustainable, the DH Course Registry is embedded within the research infrastructure as a joint DARIAH-EU / CLARIN-ERIC Working Group<sup>22</sup> and there is an ongoing project to improve sustainability through usability<sup>23</sup>. The entire registry is based on Open Source software<sup>24</sup> and is hosted by the Austrian Academy of Sciences.

Lecturers or program administrators can contribute to the platform by adding metadata of their course or programs after prior registration on the platform.

## **dariahTeach**

dariahTeach<sup>25</sup> started as an Erasmus+ project supported by a European Union Strategic Partnership Grant [3]. dariahTeach is an Open Source, multilingual, community-driven platform for high quality teaching and training materials for the digital arts and humanities that was launched in March 2017. The dariahTeach platform is based on Open Source Software (the underlying software is Moodle) and the teaching materials are available as open educational resources. In the context of the project, several teaching materials, with subjects ranging from an Introduction to the Digital Humanities in general to topics like Text Encoding and the TEI and Digital Scholarly Edition, have been created and are freely available to the public and for re-use at educational institutions. In terms of student effort, the teaching materials are divided into two categories: first, courses, which count for an equivalent of 5 or 10 ECTS points or workshops with no formal ECTS value, which can range

<sup>22</sup> <https://www.dariah.eu/activities/working-groups/dh-course-registry/>

<sup>23</sup> Digital Humanities Course Registry Sustain – Improving Sustainability through Usability (DARIAH Theme Call 2018/2019 funding) <https://www.dariah.eu/2018/11/05/dariah-theme-funding-call-2018-2019-meet-the-winning-projects/>

<sup>24</sup> Available on <https://github.com/hashmich/DH-Registry>

<sup>25</sup> <https://teach.dariah.eu/>

from several hours up to 2 days. Focussing on the aspect of a research infrastructure as a network of expertise, members of the community are encouraged to contribute to *dariahTeach*: by creating their own modules, peer-reviewing material provided by their colleagues or translating existing material. In 2017 the *dariahTeach* platform had 5735 visits and in 2018 (until 10 November 2018) 5401 visits, comprising 4290 visits from Europe, 614 from North America, 319 from Asia, and the remainder from South America, Oceania and Africa.

After the end of the funding period, *dariahTeach* was transformed into a DARIAH Working Group in order to maintain the platform hosted at the Austrian Academy of Sciences and to guarantee long-term sustainability. Furthermore, other funded projects (e.g. the European project IGNITE<sup>26</sup>, or the Visegradfund project Training Digital Scholars: Knowledge Exchange between V4 and Austria) will create more teaching materials for specific DH topics.

Lecturer and research can contribute in different forms to the #*dariahTeach* platform: They can contribute to the already existing course ‘Introduction to Digital Humanities’ with videos or with multilingual subtitles to existing videos. The other possibility is to create a whole course or workshop<sup>27</sup>.

## ACKNOWLEDGEMENTS

This contribution was supported by CLARIN ERIC and DARIAH-EU.

## REFERENCES

- [1] Clarin + D5.2 Operational course and education material registry. 2017. [https://office.clarin.eu/v/CE-2017-0986-CLARINPLUS-D5\\_2.pdf](https://office.clarin.eu/v/CE-2017-0986-CLARINPLUS-D5_2.pdf)
- [2] ESFRI ROADMAP 2019. Strategy Report on Research Infrastructures. Part 1. 2018. <http://roadmap2018.esfri.eu/media/1048/rm2018-part1-20.pdf>
- [3] Schreibman, S., Benardou, A., Clivaz, C., Durco, M., Huang, M., Papaki, E., Scagliola, S., Tasovac, T., Wissik, T. (2016). #*dariahTeach*: online teaching, MOOCs and beyond. In *Digital Humanities 2016: Conference Abstracts*. Jagiellonian University & Pedagogical University, Kraków, pp. 354-356. <http://dh2016.adho.org/abstracts/292>

---

<sup>26</sup> <https://ignite.acdh.oeaw.ac.at/>

<sup>27</sup> For more information regarding contributing to #*dariahTeach* please contact the team via e-mail [dariahteach@gmail.com](mailto:dariahteach@gmail.com)

# (Ri)progettare un master online: l'esperienza del Consorzio ICoN

Paola Andriani, Maria Chiara Tamani, Francesca Tolaini

Consorzio ICoN -Italian Culture on the Net

[andriani@italicon.it](mailto:andriani@italicon.it), [tamani@italicon.it](mailto:tamani@italicon.it), [tolaini@italicon.it](mailto:tolaini@italicon.it)

## ABSTRACT.

Questo contributo presenta l'esperienza di Web content e instructional design del Consorzio ICoN per il master in Didattica della lingua e della letteratura italiana. Progettato dalle Università per Stranieri di Siena e di Perugia, è erogato in modalità blended dall'Università per Stranieri di Siena; il Consorzio ICoN cura e gestisce la didattica online. Il master è stato pensato sin dalla prima edizione (2009) in formato digitale per un pubblico di corsisti residenti in tutto il mondo: il suo scopo è la formazione di docenti di italiano L2 all'estero. L'intervento si incentra sui criteri adottati da ICoN per l'aggiornamento dell'ambiente di apprendimento (2016): una progettazione user-based, orientata alla possibilità di ulteriori integrazioni e ad una maggiore flessibilità d'uso. Questo master presenta una struttura complessa, caratterizzata da due percorsi paralleli di apprendimento (linguistica educativa e didattica della letteratura) che sono didatticamente e cronologicamente suddivisi in due fasi (basi teoriche ed aspetti pratici) e alternati a una settimana in presenza, prove di verifica e tirocinio, per concludersi con la discussione della tesi. Curare l'edizione online di un processo di apprendimento complesso e di alto livello come quello di un master universitario è un'operazione delicata, che richiede competenza da parte degli esperti di dominio. Una (ri)progettazione lungimirante permette: di mettere gli utenti al centro del processo formativo, ai corsisti di avere i materiali sempre reperibili, di conciliare le attività con i tempi di studio e lavoro, di fruire del percorso formativo indipendentemente dai propri mezzi hardware e software, ai tutor di adattare la didattica sui propri studenti e seguirli in modalità asincrona, agli esperti del settore di avere un sistema flessibile da poter mantenere e aggiornare facilmente. Nell'esperienza di ICoN, lo studio online ha inoltre una forte interconnessione sociale e motivazionale poiché unisce persone con lo stesso obiettivo formativo, superando i limiti geografici.

## PAROLE CHIAVE

e-learning, instructional design, learning environment, learning design, moodle

## INTRODUZIONE

Questo contributo presenta l'esperienza di Web content e instructional design del Consorzio ICoN per il master in Didattica della lingua e della letteratura italiana<sup>28</sup>.

Il master è pensato per un pubblico di corsisti residenti in tutto il mondo e il suo scopo è la formazione specialistica di docenti di italiano L2 all'estero. Erogato in modalità blended dall'Università per Stranieri di Siena, concentra l'attività didattica prevalentemente online e quindi è stato progettato, sin dalla prima edizione (2009), in formato digitale.

Il nostro intervento si incentra sui criteri adottati da ICoN per l'aggiornamento dell'ambiente di apprendimento della parte online, ristrutturato per la sesta edizione (2016/2017): una

---

<sup>28</sup> <http://www.didattica.icon-master.it> <ultima visita 10 novembre 2018>.

progettazione user-based, orientata alla possibilità di ulteriori integrazioni e ad una maggiore flessibilità d'uso.

Con questo contributo vogliamo condividere quali vantaggi comporta una progettazione lungimirante dei materiali, e quali potenzialità può “liberare” il formato digitale all'interno di un processo di apprendimento complesso e di alto livello come quello di un master universitario.

## **PREMESSA**

ICoN (Italian Culture on the Net) è un consorzio di università italiane nato nel 1999 con lo scopo di promuovere e diffondere la lingua e la cultura italiana nel mondo attraverso l'e-learning e le nuove tecnologie. I corsi ICoN, realizzati in collaborazione con le università consorziate, sono strettamente collegati alle discipline umanistiche: un corso di laurea triennale in lingua e cultura italiana, un master in traduzione specialistica dall'inglese all'italiano, un master in didattica della lingua e della letteratura italiana, corsi di formazione per docenti di italiano all'estero e un ampio spettro di corsi di lingua italiana in tutorato e in autoapprendimento.

Tutti i progetti attivi del Consorzio hanno attraversato diverse fasi di aggiornamento tecnologico, ognuno progettato sulla base dell'evoluzione delle risorse e dei *know-how* in materia di *Web Based Learning* (WBL). Nel caso del master in *Didattica della lingua e della letteratura italiana*, l'aggiornamento è stato realizzato grazie ad un incrocio di competenze tra progettisti didattici, redattori e informatici del Consorzio e basato sulle necessità nate dall'esperienza dei tutor, coinvolti direttamente nel lavoro.

Abbiamo quindi scelto di presentare questo master perché esemplificativo del *modus operandi* del Consorzio ICoN: progettazione didattica collaborativa in formato digitale.

## **IL MASTER IN DIDATTICA DELLA LINGUA E DELLA LETTERATURA ITALIANA**

Progettato dalle Università per Stranieri di Siena e di Perugia, il master è erogato dall'Università per Stranieri di Siena. Il Consorzio ICoN cura e gestisce la didattica online (studio, attività didattiche e prove di verifica intermedie), mentre l'Università di Siena gestisce, oltre agli aspetti di segreteria, la settimana introduttiva in presenza, il tirocinio e la discussione della tesi finale.

Questo master presenta una struttura complessa, caratterizzata da due percorsi paralleli di apprendimento: uno destinato alla linguistica e alla didattica della lingua italiana L2 e l'altro alla didattica della letteratura italiana contemporanea. I due percorsi sono didatticamente e cronologicamente suddivisi in due fasi (basi teoriche ed aspetti pratici) e alternati a un momento in presenza, a prove di verifica e a un tirocinio nel paese di residenza del candidato, per concludersi con la discussione della tesi.

La prima edizione è stata erogata nel 2009 su una piattaforma che unificava l'attività didattica e gli aspetti di segreteria. Il sistema informativo alla base consisteva in un sito strutturato in HTML e poggiava su un'architettura server Windows, costituita da un database mSQL e dall'impiego di tecnologia ASP (*Active Server Pages*) per la creazione di pagine dinamiche. Quest'impostazione è stata ereditata dalla struttura dati del corso di laurea (2002) e permetteva alla piattaforma di sfruttare tutte le potenzialità dell'ipertesto e delle risorse multimediali disponibili all'epoca: pagine HTML, immagini, audio, video, PDF, tutto a bassa definizione per minimizzare l'effetto del *digital divide* tra paesi di residenza degli studenti.

Le attività svolte in piattaforma consistevano nello studio dei materiali, in esercizi online, e in una serie di attività di tutorato completamente basato sull'interazione con i corsisti all'interno di forum.

L'architettura è stata ristrutturata dal Consorzio nel 2016 con lo scopo di avere un sistema informativo unificato che gestisse ogni iniziativa didattica con una logica modulare. In questa nuova logica, la componente didattica è nettamente distinta da quella informativa e gestionale. Allo stato attuale è possibile implementare più piattaforme didattiche, anche differenti tra loro, sullo stesso sistema informativo. La possibilità di integrare ulteriori componenti, come è stato fatto per il nuovo sistema di tracciamento dell'attività degli studenti (LRS) e per un *repository* centralizzato di materiali didattici, garantisce la flessibilità necessaria ad un sistema in evoluzione. Al momento della ristrutturazione abbiamo riprogettato l'ambiente di apprendimento del master scegliendo di personalizzare l'LMS (*Learning Management System*) *open source* Moodle.

Per la riprogettazione del master abbiamo seguito le logiche dell'*instructional design*, integrando gli oggetti di apprendimento già presenti (testi, esercizi) con nuovi oggetti, individuati o creati in base a criteri didattici e tecnici (obiettivi formativi, tipo di risorsa o di attività, loro usabilità e sostenibilità), e assemblandoli in un percorso basato sulle necessità dell'apprendente. Abbiamo quindi integrato soluzioni proprietarie di Moodle con funzionalità sviluppate dalla *community* (descritte di seguito) e con soluzioni personalizzate (grafica, personalizzazione del foglio di stampa dei libri). Questo ha segnato un'evoluzione interessante rispetto al passato quando l'ambiente di apprendimento del master, per esempio, era il medesimo per tutti gli utenti e non consentiva percorsi personalizzati.

## **PENSARE IN DIGITALE**

L'aggiornamento del master si è svolto su due livelli: i materiali didattici e l'ambiente di apprendimento. I materiali didattici sono stati aggiornati solo dal punto di vista dei contenuti: non sono stati necessari ulteriori interventi poiché non si sono verificati problemi di compatibilità nel nuovo ambiente, né dal punto di vista tecnico né da quello di coerenza generale. I materiali, infatti, erano stati progettati da subito come elementi riutilizzabili di un unico grande testo digitale. Più articolati sono stati i criteri di progettazione della nuova piattaforma.

*La chiarezza e la semplicità d'uso.* La struttura di questo master viene esplicitata sia dalla resa grafica, progettata *ad hoc*, che dalle impostazioni di visibilità di sezioni e sottosezioni. La piattaforma rende visibili i contenuti necessari al proseguimento dell'esperienza in base alle date impostate all'inizio della corrente edizione del master, evitando di creare confusione e spaesamento. La grafica, inoltre, è stata studiata con cura per indirizzare l'attenzione dell'utente sul percorso logico da seguire, e mantenere l'attenzione sui materiali di studio. L'area dedicata allo studio dei testi, per esempio, è stata distinta graficamente dal resto del percorso didattico: escludendo gli elementi di disturbo e variando colore di background e carattere è stato creato un ambiente dedicato alla lettura e alla concentrazione.

*La reperibilità dei materiali.* Affinché lo studente abbia a disposizione i materiali di studio necessari durante la frequenza del master, all'interno delle aree in cui si svolge l'attività didattica, queste sono richiamate (e rese visibili) secondo una logica cronologica e tematica. Inoltre per rendere i materiali di studio sempre accessibili, sono state realizzate delle sezioni fisse (primi passi, scaffale digitale) contenenti tutte le risorse richiamate all'interno del master. In questo modo lo studente può reperire i materiali di studio che gli occorrono anche seguendo una logica personale.

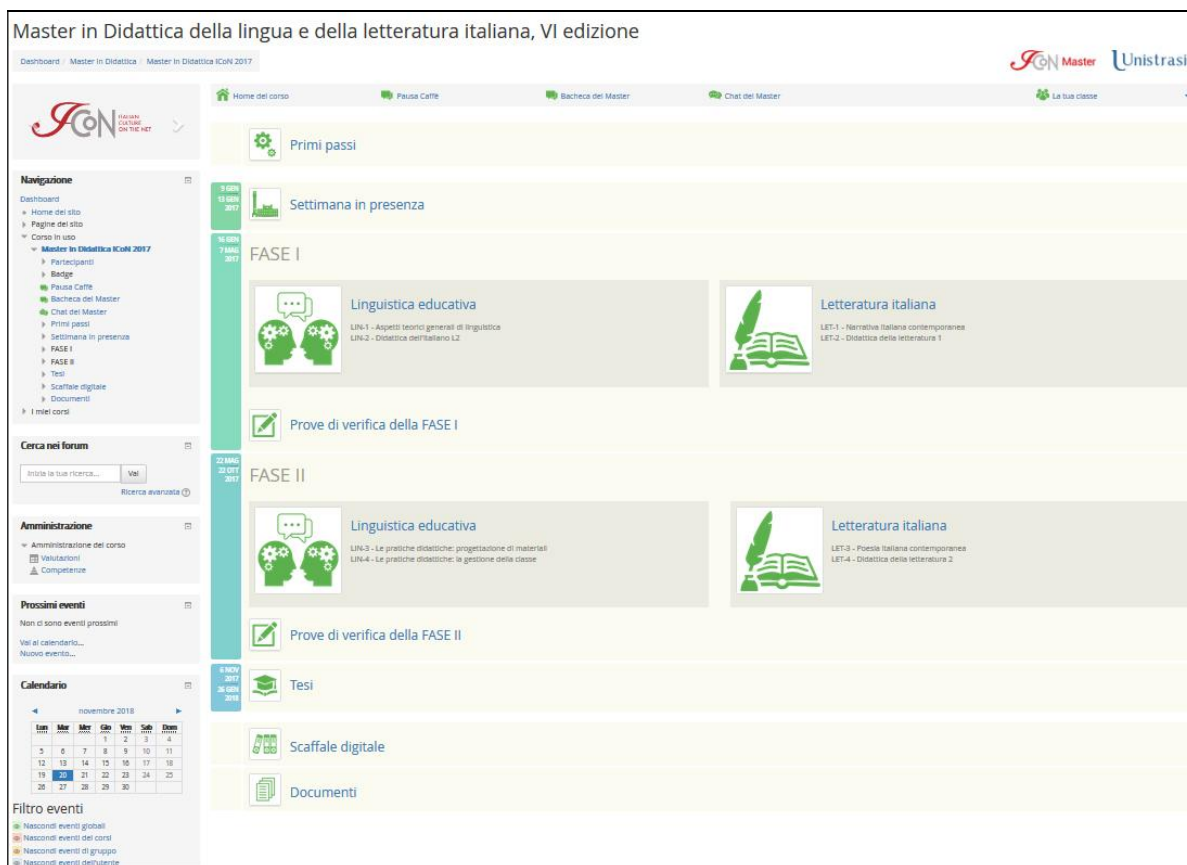


Figura 1 - La home page del master: visualizzazione utente

*La flessibilità del percorso didattico.* È stata predisposta la possibilità di accedere a un percorso personalizzato definito visivamente e limitato alle aree didattiche previste dal percorso di studio. Questo risultato è stato ottenuto mediante una personalizzazione dei ruoli e dei gruppi. I ruoli hanno il vantaggio sia di distinguere in maniera trasparente la tipologia di utente che di diversificare gli spazi a disposizione: i tutor, pur condividendo lo stesso ambiente dei corsisti, hanno accesso a un'area riservata per avviare dialoghi interni (tra loro o con lo staff di ICoN) e per consultare della documentazione specifica. Per consentire questa visualizzazione abbiamo fatto ricorso a un *plug-in* sviluppato dalla *community* di Moodle, che offre la possibilità di limitare l'accesso alle sezioni in base al ruolo dell'utente (*Availability conditions: Restriction by course role*<sup>29</sup>). I gruppi offrono un ulteriore livello di dettaglio: non solo tutor e corsisti, ma anche coordinatori della didattica e utenti associati a uno solo dei due percorsi online (linguistica oppure letteratura) che, per esempio, non hanno accesso alla sezione dedicata allo svolgimento della tesi finale.

*L'impiego di didattica asincrona.* Per agevolare l'utente, le attività sono progettate per essere seguite e svolte senza il vincolo di orari (importantissimo, nel nostro caso, per utenti che hanno fusi orari diversi). Esempi di queste attività sono esercitazioni scritte singole, strumenti di scrittura condivisa, discussioni sui forum.

<sup>29</sup> [https://moodle.org/plugins/availability\\_role](https://moodle.org/plugins/availability_role) <ultima visita 20 dicembre 2018>.

**Figura 2 - Sottosezione dedicata alle esercitazioni del percorso di Letteratura italiana, Fase II.**

Una maggiore varietà di attività. I tutor hanno a disposizione un numero maggiore di attività da proporre, siano esse collaborative (forum, wiki, database, workshop) o individuali (quiz, compito), che possono gestire in autonomia o in collaborazione con la redazione. Le attività possono inoltre essere riprogettate ad ogni edizione, seguendo le necessità della classe. Il ricorso ad un management system orientato alla didattica come Moodle offre inoltre la possibilità di integrare nuovi plug-in sviluppati dalla community o in autonomia. Per esempio, l'attività "Dialogo con il tutor" della sezione "Tesi" sostituisce lo scambio che altrimenti avverrebbe extra-piattaforma. Per questo abbiamo selezionato un plug-in open source che consente lo sviluppo di conversazioni private all'interno del corso (Activities: Dialogue<sup>30</sup>).

<sup>30</sup> [https://moodle.org/plugins/mod\\_dialogue](https://moodle.org/plugins/mod_dialogue) <ultima visita 20 dicembre 2018>.

The screenshot shows the ICoN Master LMS interface. At the top, the page title is 'ICoN Master' and the user is logged in as 'Amministratore ICoN'. The breadcrumb trail indicates the current location: 'Dashboard / Corsi / Master in Didattica / Master in Didattica ICoN 2018 / Tesi / Tesi - Dialogo col tutor'. The navigation menu on the left includes 'Dashboard', 'Home del sito', 'Pagine del sito', 'Corso in uso', 'Master in Didattica ICoN 2018', 'Partecipanti', 'Badge', 'Pausa Caffè', 'Bacheca del Master', 'Chat del Master', 'Primi passi', 'Settimana in presenza', 'FASE I', 'FASE II', 'Tesi', 'Tesi - Dialogo col tutor', 'Forum per tutor e testisti', 'Materiali generali tesisti', 'Indicazioni per la stesura della tesi', 'Linee guida antiplagio', 'Come ridurre la dimensione dei pdf', 'Tirocinio - nascosta', 'Scaffale digitale', 'Documenti', and 'Corsi'. The main content area shows a chat window titled 'Tesi' with a 'Torna a 'Tesi'' button and an 'Aperte' status. The chat history shows a message from Marianna Marrucci dated 27 November (9:20 pm) and a reply from Elena dated 27 November (9:45 pm). The chat participants are listed as '2 partecipanti' with avatars for Elena and Marianna Marrucci.

Figura 3 - Una conversazione tra tutor e corsista all'interno dell'attività dialogo.

*Il potenziamento degli operatori didattici.* Il ricorso ad un LMS dotato di un sistema di tracciamento ha permesso a tutor e coordinatori delle attività didattiche di monitorare la navigazione e lo svolgimento delle attività degli utenti.

*L'interconnessione sociale.* Lo spazio sociale di apprendimento è stato ampliato ricorrendo ad accorgimenti quali una rappresentazione grafica della classe virtuale e il collegamento ai social network, ormai indispensabile nell'era del Web sociale.

*L'accessibilità dello strumento.* La piattaforma è fruibile tramite tutti i browser di navigazione e sui dispositivi mobili grazie alle soluzioni *responsive* già integrate nell'LMS e nel *framework* utilizzato per le soluzioni grafiche (Bootstrap).

Uno degli obiettivi della prossima edizione è quello di integrare direttamente in piattaforma un questionario anonimo specifico per il master, in modo da ricevere un feedback completo sulla loro esperienza didattica digitale.

## CONCLUSIONI

In ultima analisi, i vantaggi portati dalla riprogettazione di questo master sono molteplici: (I) gli utenti coinvolti nel processo formativo si trovano al centro di un percorso ritagliato su misura, visivamente chiaro; (II) i corsisti hanno a disposizione materiali di apprendimento sempre reperibili in ogni parte del mondo; (III) le attività possono essere svolte conciliando i propri tempi di studio e di lavoro; (IV) il percorso è fruibile indipendentemente dall'hardware e dai browser disponibili; (V) il tutor può seguire l'apprendente in modalità asincrona senza rinunciare alla differenziazione delle attività; (VI) la flessibilità del sistema consente agli esperti di settore di mantenere e aggiornare facilmente il master.

Nell'esperienza di ICoN, lo studio online ha inoltre una forte interconnessione sociale e motivazionale poiché unisce persone con lo stesso obiettivo formativo, superando i limiti



geografici: ogni studente può giovare dell'accesso a materie altrimenti precluse, e arricchirsi personalmente e culturalmente.

## RINGRAZIAMENTI

La riprogettazione del master è stata realizzata con la collaborazione dello staff ICoN (in particolare di Elisa Bianchi, Marco Infortugno e Maria Teresa Lupia) e delle tutor Cristina Gavagnin e Marianna Marrucci.

## BIBLIOGRAFIA

- [1] Arzilli, C., Bianchi, E., Infortugno M. In corso di stampa. *A new Moodle-based learning environment for Italian as a second language*. In: Atti del convegno *Innovation in language learning, XI*, Bologna: Filodritto editore <<https://conference.pixel-online.net/ICT4LL/files/ict4ll/ed0011/FP/1223-ICL3416-FP-ICT4LL11.pdf>>
- [2] Bonaiuti, G. 2006. *E-learning 2.0. Il futuro dell'apprendimento in rete, tra formale e informale*. Trento: Erickson.
- [3] Calvani, A., Menichetti, L. 2005. *Come fare un progetto didattico. Gli errori da evitare*. Roma: Carocci.
- [4] Giacomantonio, M. 2007. *Learning object. Progettazione dei contenuti didattici per l'e-learning*. Roma: Carocci.
- [5] ICoN - Italian Culture on the Net. 2014. *Statuto del Consorzio ICoN* <<http://www.italicon.education/sites/default/files/statuto-icon-in-vigore-da-2014.pdf>>.
- [6] Laici, C. 2007. *Nuovi ambienti di apprendimento per l'e-learning*. Perugia: Morlacchi Editore.
- [7] Paternò, F. 2001. *Interazione Uomo-Computer: Un'Introduzione*. In: Mondo Digitale, N.4, Dicembre 2004 <<http://giove.isti.cnr.it/Corso/mondo-digitale.pdf>>.
- [8] Pedroni, M. 2006. *E-learning e rappresentazione della conoscenza*. Ferrara: Telecom Project.
- [9] Roncaglia, G. 2013. *L'uso degli e-book nella formazione e nella didattica: tre riflessioni sugli ebook di testo e sul loro ruolo*. In: dispense del corso "E-book e futuro del libro" – Università della Tuscia. <<http://www.bookinprogress.org/file/documenti/seminario/Approfondimento%20-%20gli%20e-book%20a%20scuola.pdf>> (visitato il 05/08/2016).
- [10] Tamani, M. C. 2017. *La riusabilità dei contenuti didattici ICoN*. Tesi di laurea magistrale in Informatica Umanistica. Pisa: Università degli Studi di Pisa.
- [11] Trentin, G. 2008. *La sostenibilità didattico-formativa dell'e-learning. Social networking e apprendimento attivo*. Milano: Franco Angeli.
- [12] Varanini, F. 2013., *La letteratura ai tempi del codice digitale e i nuovi incunaboli*. Università di Pisa, 23 ottobre 2013. <<http://www.francescovaranini.it/2013/09/la-letteratura-ai-tempi-del-codice-digitale-e-i-nuovi-incunaboli-universita-di-pisa-23-ottobre-2013/>> (visitato il 12/05/2017).
- [13] Villarini, A. 2018. *Metodi, strumenti e ambienti per la didattica dell'italiano con il Web*. Pisa: Consorzio ICoN.

# Per un'edizione *online* dell'Epistolario di Alcide De Gasperi. Criteri di digitalizzazione, schedatura, registrazione ed edizione di lettere del Novecento

Stefano Malfatti

Università degli Studi di Trento - Edizione Nazionale dell'Epistolario di Alcide De Gasperi  
Via T. Gar, 14 - via S. Croce, 77 - 38122 Trento, s.malfatti@epistolariodegasperi.it

## ABSTRACT

L'Edizione nazionale dell'Epistolario di Alcide De Gasperi, completamente digitale, intende raccogliere, digitalizzare, schedare e presentare la corrispondenza dello statista reperita in archivi italiani e stranieri. Gli esiti dell'iniziativa editoriale, che si contraddistingue per l'essere frutto del lavoro multidisciplinare di storici, archivisti, filologi e specialisti in Digital Humanities, saranno pubblicati nel febbraio 2019 su una piattaforma web *open access* ([www.epistolariodegasperi.it](http://www.epistolariodegasperi.it)). Diversamente dalle tradizionali edizioni cartacee, l'edizione *online* consentirà di ricavare numerose informazioni dai documenti, visto l'elevato numero di metadati raccolti durante la schedatura. Immagini, metadati ed edizione vengono infatti caricati dai collaboratori dell'Edizione nel software 'LETTERE' ('LETters Transcription Environment for Research'), sviluppato dal gruppo DH-FBK. Il programma permette di raccogliere metadati relativi ai caratteri intrinseci (titolo, data topica e cronica, mittente, destinatario, tipologia del documento, lingua ecc.) e ai caratteri estrinseci (buste o intestazioni, descrizione fisica, stato di conservazione ecc.) del documento, oltre alle informazioni relative alla segnatura archivistica e all'eventuale presenza di edizioni. Prima della trascrizione, è prevista la registrazione dell'abstract, necessario per contestualizzare e riassumere il contenuto della corrispondenza. Per quanto concerne l'edizione critica, in linea generale è prevista una trascrizione fedele del testo, senza interventi mediante correzioni; eventuali errori possono essere segnalati attraverso il sistema automatico di inserimento delle note. È previsto infine un secondo ordine di note, di tipo storico, o di commento, utili a chiarire il contenuto del documento. In ultimo, per quanto riguarda il *database* di consultazione dell'Edizione, sono previste due maschere di ricerca (semplice e avanzata), nonché quattro 'percorsi tematici' (per fasi biografiche dello statista, per archivi, per ambito geografico [a partire dalla data topica del documento] e per corrispondenti), che avranno lo scopo di aiutare il ricercatore ad individuare con facilità le lettere d'interesse.

## PAROLE CHIAVE

Edizione critica *online*, Alcide De Gasperi, storia contemporanea, metadati, trascrizione, digitalizzazione, lettere, corrispondenza, *database*, *tool*.

## INTRODUZIONE

L'istituzione dell'Edizione nazionale dell'epistolario di Alcide De Gasperi risponde alla necessità di raccogliere e valorizzare il grande patrimonio di scritti epistolari prodotto dallo statista e dai suoi corrispondenti italiani e stranieri. Da un lato, infatti, l'Opera intende proseguire e completare il lavoro di edizione degli *Scritti e discorsi politici* degasperiani, dall'altro vuole colmare il parziale vuoto di fonti epistolari edite: si deve infatti constatare come soltanto una minima porzione della corrispondenza degasperiana sia stata, fino ad oggi, oggetto di edizione. Quanto mai importante si rivela dunque l'iniziativa editoriale, tanto più perché la documentazione si trova oggi disseminata in numerosissimi archivi pubblici e

privati. La pubblicazione *online* non solo garantisce l'accesso immediato a fonti rimaste fino ad oggi quasi del tutto sconosciute, ma si configura come strumento essenziale per conoscere ancor più profondamente l'operato di De Gasperi e dei suoi corrispondenti. Al contempo, considerata anche la mole documentaria da prendere in considerazione, si rende necessaria la collaborazione di numerosi ricercatori, a formare così un *team* multidisciplinare di storici, archivisti, filologi e informatici, ciascuno dei quali può mettere a frutto le proprie competenze. L'edizione del carteggio degasperiano diviene dunque, al contempo, strumento di conoscenza e prodotto scientifico pienamente riconosciuto (anche grazie all'associazione di codici ISBN e DOI). Non meno importanti sono le plurime occasioni di confronto e collaborazione fra il settore più prettamente 'umanistico' degli storici, degli archivisti, dei filologi che hanno proposto i criteri per la raccolta e l'edizione della corrispondenza e il settore 'digitale' degli informatici che hanno sviluppato i *software* per la raccolta e la pubblicazione delle lettere.

### a. IL PROGETTO

Ogni collaboratore dell'Edizione nazionale si reca presso l'archivio per raccogliere i documenti; ove possibile, previa stipulazione di un accordo con l'istituto di conservazione, viene acquisita l'immagine digitale delle lettere. La digitalizzazione può essere affidata allo stesso archivio oppure fatta dal collaboratore, purché siano rispettati gli standard minimi previsti per formato (.jpg) e risoluzione (non meno di 300 dpi). Affinché tutti i collaboratori dell'Edizione nazionale svolgano il lavoro di ricerca e acquisizione della corrispondenza correttamente, è stato predisposto un vademecum<sup>31</sup> contenente tutte le informazioni necessarie, oltre ai criteri di edizione per la trascrizione delle lettere.

Si deve al gruppo DH-FBK (Giovanni Moretti, Rachele Sprugnoli, Sara Tonelli) la costruzione del *tool* "LETTERE" (LETters Transcription Environment for REsearch), software *stand-alone*, sviluppato per la descrizione e l'edizione di corrispondenza. Il ricercatore, dopo aver scaricato l'applicazione sul proprio computer, e dopo aver effettuato il login, può iniziare la registrazione di metadati, la trascrizione e il regesto del documento, corredandoli poi con l'immagine digitale. Il lavoro si salva immediatamente in locale e, in presenza di connessione internet, sul server dell'Edizione nazionale.

Conclusa la fase di controllo sul documento da parte del collaboratore, i dati vengono inviati alla revisione, che è scandita in due distinte fasi: anzitutto è svolto un controllo da parte del supervisore; in seguito spetta alla Redazione svolgere un'ultima revisione per verificare il rispetto dei criteri editoriali. La lettera edita può quindi essere pubblicata *online* sul *database* di consultazione.

Fra i principali metadati da raccogliere (Fig. 1) vi sono: mittente e destinatario, per i quali può essere specificato anche il genere (maschile o femminile), la data cronica (che, con i dati precedenti, va a comporre il titolo del documento), la data topica, la tipologia del documento (lettera, telegramma, cartolina ecc). Dal punto di vista contenutistico va determinato se si è in presenza di uno scritto di natura personale o istituzionale, oltre al tema/temi trattati, da scegliere fra i nove forniti nel menù. Fra i caratteri intrinseci si ha inoltre la lingua di redazione. In ultimo, va chiarito se si tratti di un originale, di una copia o di una minuta.

Fra i caratteri estrinseci vanno invece rilevati: la presenza, o meno, di una busta, di carta intestata, di firme autografe e di allegati. Per ciascun campo si può fornire una descrizione. Per ogni documento vanno inoltre indicati: il numero delle carte, le misure in millimetri, la tecnica di scrittura (manoscritta o dattiloscritta) e lo stato di conservazione. Necessari alla corretta identificazione del documento sono la segnatura archivistica ed, eventualmente, i riferimenti ad altre edizioni. Precede l'ampio spazio destinato a raccogliere la trascrizione, il

---

<sup>31</sup> Reperibile all'indirizzo <http://epistolariodegasperi.it/risorse.jsp> (ultimo accesso: 30/12/2018).

campo abstract, necessario per contestualizzare e riassumere il contenuto del documento (Fig. 2).

Per quanto riguarda i criteri di edizione, in linea generale, si è scelto di trascrivere il testo con assoluta fedeltà, rispettando ortografia e punteggiatura; eventuali errori possono eventualmente essere segnalati producendo, con l'ausilio dell'apposito menù nel *tool*, una nota critico-testuale (atta a segnalare refusi, cancellature, lacune per guasto meccanico ecc.). Le note storiche o di commento chiariscono il contenuto del documento con integrazioni e informazioni che ne consentano una migliore comprensione: ad esempio forniscono il nome completo di un personaggio citato nel testo in modo incompleto o attraverso pseudonimi; forniscono spiegazioni in relazione a fatti o eventi non citati nella lettera in modo sufficientemente chiaro ecc.

Inserimento Dati dati necessari alla schedatura documento

**TITOLO**  
Da Alcide De Gasperi A Giovanni Ciccolini , 14/02/1925

**Mittente/i**  
 Nome: Alcide  M  F  **Dato provvisorio:**   
 Luogo: Trento  **Dato provvisorio:**   
 Cognome: De Gasperi

**Destinatario/i**  
 Nome: Giovanni  M  F  **Dato provvisorio:**   
 Luogo: NA  **Dato provvisorio:**   
 Cognome - Destinatario collettivo/altra denominazione: Ciccolini

**Data del documento:** 14/02/1925  **Dato provvisorio:**   
**Data di acquisizione:** 11/06/2018

**Tipologia:** Lettera  
**Tema:** Tema  
 Vita Privata  
 Politica Nazionale  
 Politica Internazionale  
 Politica Locale

**Tipo di scritto:** Personale  
**Esemplare:** Originale  
**Diritti di utilizzo:** Dominio pubblico / nessuna restrizione d'uso

**Lingua:** Italian  **Busta:**  **Descrizione Busta:**

**Carta Intestata:**  **Intestazione:** Intestazione  
**Numero Carte:** 1 **Larghezza:** 210 mm **Altezza:** 297 mm **Tecnica di scrittura:** Manoscritta

**Descrizione fisica:** Lettera vergata su carta di colore bianco, con inchiostro di colore bruno scuro  **Firma autografa:**  **Firma:** Degasperì

**Stato di conservazione:** Buono

**Deposito/Archivio:** Deposito/Archivio **Sigla:** Codice Arch **Segnatura:** Segnatura

**Edita:** Edita

**Commento sui metadati:**

Fig. 1. Sezione superiore della schermata del tool 'LETTERE' con i metadati da compilare.

Abstract (massimo 600 caratteri)

De Gasperi scrive a Giovanni Ciccolini, quale presidente della Società laureati cattolici, per ringraziarlo della solidarietà espressa nei suoi confronti in occasione della recente campagna denigratoria fascista.

Allegati

Carica Immagine

Trascrizione:

Trento, 14.2.25

Caro Ciccolini,  
 a te quale presidente della società laureati cattolici, un vivo ringraziamento per l'atto di solidarietà fraterna. Ne sono lusingato più che di qualsiasi altra adesione.  
 La stima degli uomini onesti sana le ferite prodotte dai dardi avvelenati della denigrazione. Gli avversari mi colpiscono nell'orgoglio della mia vita, cioè nel sentimento di italianità. Ma mi conforta il pensiero vostro che è quello di <sup>chi</sup> può testimoniare la verità per averla vissuta. Ti prego di ringraziare sentitissimamente tutti coloro che hanno contribuito alla manifestazione di simpatia per me e di assicurarmi della mia gratitudine. A te in particolare un abbraccio affettuoso

tuo Degasper

Postille

Rimuovi Pagina « < > » Aggiungi Pagina

Salva

Lettere - Digital Humanities FBK

Fig. 2. Sezione inferiore della schermata del tool 'LETTERE'; sezione dedicata all'abstract, all'immagine digitale e alla trascrizione.

## CONCLUSIONI E SVILUPPI FUTURI

Attualmente è in corso la raccolta della corrispondenza degasperiana da includere nell'epistolario *online*; entro la fine del 2018 si concluderà la prima fase del progetto con la verifica del lavoro finora svolto, mentre è prevista nel febbraio 2019 la pubblicazione *online* del primo *corpus* di lettere, che consta attualmente in circa 1500 documenti.

La piattaforma di consultazione prevederà due maschere di ricerca tradizionali (semplice e avanzata); la prima consentirà di svolgere la ricerca attraverso pochi campi: testo libero (con ricerca fra tutti i metadati, nella trascrizione e nelle note); per nome (con ricerca nei campi mittente e destinatario) e per date topica e cronica. La maschera avanzata permetterà invece la ricerca attraverso più campi: per tipologia, per tema, per tecnica di scrittura ecc. Sono inoltre previsti quattro percorsi tematici che "aiuteranno" l'utente nella ricerca nel *database*: per fasi biografiche (definite sulla base di cesure cronologiche) e, come ulteriori elementi di

affinamento della ricerca, per ambiti tematici e per cronologia; per archivi, per cui all'interno di una mappa navigabile saranno selezionabili gli archivi (e poi i fondi) in cui si è reperita corrispondenza degasperiana; per ambito geografico (da individuare sempre tramite mappa navigabile), tratto dalla data topica del documento; in ultimo, per corrispondenti.

Le singole lettere edite potranno essere visualizzate con le note a vista, e dunque con tutti gli interventi atti a evidenziarne lacune, parti di testo sovrascritte ecc., oppure in modalità "pulita", cioè senza note. Le note vengono segnalate per mezzo di colori diversi, a seconda che si tratti di note di commento o note critiche al testo. Esse saranno visualizzabili sulla piattaforma attraverso i rispettivi colori (selezionando il termine o la stringa di testo annotata, sarà possibile leggere la nota) o in modo tradizionale (con lettere dell'alfabeto minuscolo per le note critiche al testo e con numeri per le note di commento/storiche), nel qual caso esse saranno disposte a piè di pagina.

Il testo del documento sarà inoltre preceduto dal titolo e dal suo abstract; tutti gli altri metadati saranno invece raccolti in specifica area, in margine alla schermata principale, ove pure troveranno spazio i dati necessari alla corretta citazione del documento dall'Edizione nazionale. In ultimo, dove opportunamente acquisita, sarà possibile visionare l'immagine digitale, filigranata, del documento.

Dopo molti mesi di preparazione il progetto dell'Edizione nazionale si appresta a prendere ufficialmente il via con la pubblicazione della prima *tranche* di lettere; miglioramenti e correttivi sono stati possibili anche in fase di caricamento delle lettere nel *software* grazie alla collaborazione dei ricercatori che hanno raccolto e trascritto la corrispondenza. La piattaforma sarà inoltre dotata in futuro di una apposita sezione destinata a raccogliere contributi scientifici inerenti alla documentazione ivi edita.

## BIBLIOGRAFIA

- [1] Carucci, P. 1987. *Il documento contemporaneo. Diplomatica e criteri di edizione*. Roma: La nuova Italia scientifica.
- [2] De Gasperi, A. 2006. *Scritti e discorsi politici di Alcide De Gasperi. Alcide De Gasperi nel Trentino asburgico*, a cura di E. Tonezzer, M. Migaran, M. Guiotto. Bologna: Il Mulino.
- [3] De Gasperi, A. 2008a. *Scritti e discorsi politici di Alcide De Gasperi. Alcide De Gasperi dal Partito popolare italiano all'esilio interno 1919-1942*, a cura di M. Bigaran, M. Cau. Bologna: Il Mulino.
- [4] De Gasperi, A. 2008b. *Scritti e discorsi politici di Alcide De Gasperi. Alcide De Gasperi e la fondazione della Democrazia cristiana, 1943-1948*, a cura di V. Capperucci, S. Lorenzini. Bologna: Il Mulino.
- [5] De Gasperi, A. 2009. *Scritti e discorsi politici di Alcide De Gasperi. Alcide De Gasperi e la stabilizzazione della Repubblica 1948-1954*, a cura di S. Lorenzini, B. Taverni. Bologna: Il Mulino.
- [6] Sprugnoli, R., Moretti, G., Tonelli, S. 2018. *Temporal Dimension in Alcide De Gasperi: Past, Present and Future in Historical Political Discourse*. In: *Book of Abstracts. Patrimoni culturali nell'era digitale. Memorie, culture umanistiche e tecnologia*, a cura di D. Spampinato. Bologna: Associazione per l'Informatica umanistica e la Cultura Digitale.
- [7] Stadler, P., Illetschko, M., Seifert, S. 2016-2017. *Towards a Model for Encoding Correspondence in the TEI: Developing and Implementing*, in *Journal of the Text Encoding Initiative* [Online], Issue 9, September 2016 - December 2017, Online since 24 September 2016, connection on 12 November 2018. URL : <http://journals.openedition.org/jtei/1433>; DOI: 10.4000/jtei.1433.





# POSTER PAPERS

# Can the “digital” reinforce a museum’s mission - what information technologies to introduce into the planned exhibition?

J. Valach<sup>1</sup>, M. Eisler<sup>1</sup> & P. Štefcová<sup>2</sup>

<sup>1</sup> Institute of Theoretical and Applied Mechanics, AS CR, Prague

<sup>2</sup> National Museum, Prague

## ABSTRACT

This paper discusses new opportunities in museum experience enhancement assisted by information technologies such as virtual and augmented realities as well as access to “big data” and multi-disciplinary data. The paper's main thesis stresses the key importance of a traditional display of artefacts and the selection of artefacts as a decisive act of interpretation imposing meaning upon a collection, while the “digital” enables individual exploration of a subject as demonstrated on discussed applications.

## KEYWORDS

Museums, exhibition, virtual reality, augmented reality, digital humanities, cuneiform tablets.

## INTRODUCTION

One of the most prominent roles museums play in our society is to provide access to the vast collection of artefacts reflecting the natural, technological and cultural world that man inhabits and builds. Museums curate and conserve their collections, and display them based on creative selections called exhibitions. An exhibition is more or less an enlightened perspective on a collection, introducing one of its innumerable meanings, and manifesting by grouping the artefacts together in a particular way. In order to exploit the role of museums in enhancing education, emphasis is placed on the comprehensiveness and self-sustainability of an exhibition. The presented paper contemplates possibilities of reaching this goal with the help of contemporary information technologies.

## THE EXHIBITION

The exhibition's primary goal is to present the results of a five-year research project called 'Analysis, description and archivation of aggregate information on properties of cultural heritage artifacts and usage of such data in restoration, conservation and research' [1-2]. It will open at the end of the project in 2020. As one of the main aims of the project is to build a comprehensive multi-disciplinary database covering Prague's cuneiform tablets collection, the exhibition will utilize data acquired by various methods and stored in the database. The exhibition will also explore opportunities for presenting the project's accumulated knowledge on its subject as well as methods (among others, data mining and visualization) of demonstrating the state of the art in digital humanities.

The second narrative line of the exhibition will be in some sense a meta-exhibition: on its rather narrow scope of displayed artefacts (mainly the collection of cuneiform tablets), it will demonstrate the succession of events starting with the tablets' creation and everyday use by the inhabitants of the ancient world, the event leading to the tablets' burial and disappearance in time, then how they were rediscovered and excavated, and finally their becoming museum items. In addition to this, museum practices such as preventive conservation and collection

curation will be demonstrated on the handling of the cuneiform collection and the methods used to acquire the presented data. The exhibition's self-awareness and self-reflection in the form of a 'heat-map' data visualization technique based on a color-coded depiction of visitors' stay duration in front of individual showcases will also be a part of the display.

## **DEVICES AND TOOLS USED TO FACILITATE THE DIDACTIC GOALS OF THE EXHIBITION**

What follows is a rather crude list of technologies meant to be discussed in detail in the final presentation.

There is a clear need for authentic objects - the four thousand year history of the tablets inspires awe and respect, and an emotional connection to the past. The availability of 3D digital models of the objects enables hands-on experience with their 3D printed replicas utilizing principles of experimental archaeology – manipulation and sensory reception of the object can be directly experienced without the mediation of a verbal description. At the same time, there is no threat to the originals safely stored in the museum repository.

As the authentic objects are kept out of reach under a glass cover, virtual reality technologies based on digital models of the object enable its manipulation, a selection of simulated illuminations, the possibility of making the object partially transparent to show its inner structure and often revealing the process with which it was made, the presence of defects, etc. The ability to rotate a cuneiform tablet's virtual model is especially important as the shadows cast by grooves are essential to the legibility of cuneiform characters.

Augmented reality – only a fraction of what is known about an exhibit can be written in the caption associated with the exhibit. Therefore, the curator has to painfully choose what is to be displayed, while the rest remains unrevealed. Nowadays, technologies and tools of augmented reality (AR) enable visitors to access all facets of information connected to the object and let the visitor discover what interests them. One can see the conflict between whatever knowledgeable storyline given by the exhibition's curator on the one hand and the wealth of individual paths to learning and discovering enabled by AR on the other hand. Layers of knowledge available for exploration through AR could be for example map data from the excavation site, a transcript of the cuneiform message, the material composition of the tablet, the wider context of other objects found at the excavation site, etc. From this point of view, the AR tools in the exhibition can be understood as an experience of an infinite expanding knowledge of the current state of the art in a narrow field of study (cuneiform tablets).

## **SUMMARY**

Digital technologies offer tools to reconcile the contradiction between a rigid frame of reference imposed by the concept of the exhibition and the individual needs of exploration of the subject. An exhibition is a decision – a contemplated selection of what and how to present. Nowadays, there is omnipresent encouragement to communicate knowledge in the form of storytelling as the most fundamental way our culture developed and was passed on to younger generations in the past. There are personal preferences behind the selection of elements used to build a story – choosing pieces that fit together in a consistent way. But there are other ways “to connect the dots” – as reality is a totality of all possible stories – and the introduction of digital technologies in exhibitions enables this simultaneity.

New opportunities enabled by information technologies bring about an expansion of the exhibition space in directions previously unthinkable. However, it is necessary to control this expansion in order to protect the message intended to be communicated in the concept of the exhibition and not to have the artefacts overshadowed by technology.

## ACKNOWLEDGEMENTS

This work is carried out within project “Analysis, description and archivation of aggregate information on properties of cultural heritage artifacts and usage of such data in restoration, conservation and research” is supported by the program of applied research and development of national and cultural identity (NAKI) of the Ministry of Culture of the Czech Republic – grant No. DG16P02M022.

## REFERENCES

- [1] Štefcová, P., Kohout, O., Polák, L., Valach, J., Zemánek, P. Comprehensive approach to preventive care of cultural heritage objects. In: Amoêda, R., ed. Heritage 2018 – Proceedings of the 6th International Conference on Heritage and Sustainable Development. Barcelos: Green Lines Institute, 2018, pp. 961-968. ISBN 9788433862617
- [2] Valach, J., Štefcová, P., Zemánek, P. A complex database for documentation of cuneiform tablet collection enabling cross-domain queries. In: Williams, D., ed. Archiving Conference, Archiving 2018 Final Program and Proceedings. Springfield (VA): Society for Imaging Science and Technology, 2018, pp. 120-123. ISBN 978-0-89208-333-6. DOI: 10.2352/issn.2168-3204.2018.1.0.26.

# Discovering Research Themes in Scientific Research: from Keyphrase Extraction to Co-occurrence Networks

Rachele Sprugnoli, Giovanni Moretti

Digital Humanities Group, Fondazione Bruno Kessler, Trento  
[rachele.sprugnoli@alumni.unitn.it](mailto:rachele.sprugnoli@alumni.unitn.it); [moretti@fbk.eu](mailto:moretti@fbk.eu)

## ABSTRACT

In this abstract we propose a method for creating co-word networks starting from keyphrases automatically extracted from the full text of scientific papers. This approach aims to overcome the limitations of the methodologies traditionally used in bibliometrics with the final goal of identifying the themes considered crucial in a research field and analyzing and visualizing its intellectual structure. In this proposal, we present the workflow and we provide a case study using the abstracts of AIUCD 2015 and AIUCD 2018 conferences.

## KEYWORDS

Text analysis, keyphrase extraction, content analysis, network analysis, bibliometrics.

## INTRODUCTION

Quantitative methods for mapping the structure and discover the trends of a research field have a long tradition in bibliometrics [1]. Most of these methods, such as co-authorship, co-citation and co-word analysis, are based on graph theory approaches that take into consideration pairwise relations between connected entities. In particular, co-word analysis uses the co-occurrence of pairs of words or phrases in a collection of papers to identify the linkages between various research themes expressed in those papers [2]: social network analysis techniques and tools are then adopted to visualize the co-occurrences and calculate standard metrics [3]. Usually, words and phrases forming co-occurrence networks are the keywords chosen by the authors during the submission phase or by professional indexers after the publication. However, this method has some limitations [4]. In case the keywords are added by professional indexers, the co-word analysis can be affected by the so-called “indexer effect” in which indexers’ views, reflected in the chosen keywords, may influence the final results. On the other hand, the use of controlled vocabularies can produce inconsistencies because different vocabularies can follow different keyword classifications. Finally, keywords chosen by authors could express more an idea with which they hope to attract the readers, than the actual content of the paper in all its facets. In addition, it is important to note that keywords are not always available: not all paper formats require them to be specified.

More recently, also topic modeling techniques [8] [9] have been used to study the development of scientific ideas [6] [7]. Nevertheless, topic modeling has some weaknesses, including the fact of not integrating linguistic information, being completely statistical, and not allowing the extraction of multi-token expressions.

For all these reasons, in this abstract, we propose another approach that takes into account the full text of a paper, automatically extracts a ranked list of keyphrases (both single words and multi-token expressions) and creates a list of edges ready to be imported in Gephi<sup>1</sup>, an open-

---

<sup>1</sup> <https://gephi.org/>

source network analysis and visualization tool [5]. The edges correspond to the co-occurring relations between keyphrases, each represented by a node in the network.

## WORKFLOW

The workflow has 3 main steps: (i) extraction of keyphrases from each paper in a corpus of scientific articles, (ii) conversion of the output of the first step into an edge list, (iii) analysis and visualization of the co-occurrence network in which the nodes are the extracted keyphrases linked to each other by an edge in case of co-occurrence within the same text.

As for the first step, in order to overcome the limitations described in the previous section and obtain more linguistically motivated and complex keywords, we use KD, Key-phrase Digger [10]. KD identifies key-concepts formed by one or more words on the basis of a list of language-specific part-of-speech patterns and combines this linguistic information with statistical information to create a weighted list of keyphrases. Furthermore, KD contains a function that recognizes and assigns a particular weight to unexpanded acronyms (eg. XML) that play a central role in scientific articles. KD can be used both through a Java implementation<sup>2</sup> and an online demo<sup>3</sup>. In both cases, the user can download the results in a tab-separated file.

An R script takes in input the output files of KD and generates a list of edges in a format suitable for being imported into Gephi as an undirected graph. The script can be customized so to include a different number of keyphrases in the graph.

Finally, the edge list is imported in Gephi: several layout algorithms and standard metrics can be chosen to check the correlation degree among keyphrases and detect communities.

## CASE STUDY

To test the procedure described in the previous section we created keyphrase co-occurrence networks of AIUCD 2015 and AIUCD 2018 abstracts [11]. The main theme of both editions was cultural heritage<sup>4</sup> but the conferences were open to other topics related to Digital Humanities. An interactive version of these networks is available in a shared folder together with the R script<sup>5</sup>.

We lemmatized the texts of all the proposals (54 abstracts for AIUCD 2015 and 67 for AIUCD 2018): English abstracts were processed with Stanford CoreNLP [12] whereas for those written in Italian we used Tint [13]. We then run KD and we translated Italian keyphrases into English. With the R script we created an edge list for each conference using the 5 keyphrases with the greater weight extracted from each abstract: this led to the creation of a network with 166 nodes and 510 edges for AIUCD 2015 and with 222 nodes and 656 edges for AIUCD 2018.

Node degree, that is the number of links that connect a node to others, was used to analyse core keyphrases: those with higher degree correspond to the topics that are more actively researched in the community. Table 1 shows an ordered list of the top 10 nodes in terms of degree in the two networks. It is interesting to note that the keyphrase representing the main theme of the conferences (*cultural heritage*) loses its absolute centrality in 2018 making room for keyphrases related to the semantic field of text analysis/encoding. While AIUCD 2018 is

---

<sup>2</sup> <http://dh.fbk.eu/technologies/kd>

<sup>3</sup> [http://celct.fbk.eu:8080/KD\\_KeyDigger/](http://celct.fbk.eu:8080/KD_KeyDigger/)

<sup>4</sup> In 2015 the title of the AIUCD conference was “Digital humanities and cultural heritage: what relationship?”; in 2018 it was “Cultural Heritage in the Digital Age. Memory, Humanities and Technologies”.

<sup>5</sup> <https://dh.fbk.eu/technologies/co-occurrence-networks-keyphrases>

more focused on the various forms of knowledge diffusion (*text, corpus, manuscript*), in 2015 the references to the institutions and infrastructures dedicated to their collection and study are more central (*library, archive, museum*). Moreover, *digital humanities* disappears (it is not a node in the AIUCD 2018 network): the field of research is now so mature that it is no longer necessary to make it explicit.

<b>AIUCD 2015</b>	<b>AIUCD 2018</b>
cultural heritage	project
digital	text
project	data
library	analysis
digital humanities	corpus
archive	digital
museum	metadata
digital technology	system
institution	manuscript
web	cultural heritage

Table 1. Top 10 nodes in terms of degree in descending order.

We also performed community detection using the Girvan-Newman clustering algorithm implemented in a Gephi plugin [14]. This algorithm identified 9 clusters in AIUCD 2015 and 13 in AIUCD 2018. The greater attention given to the syntagmatic axis, and therefore to the text, in 2018 is confirmed by the largest community (53 nodes, i.e. 24% of all nodes) detected in the AIUCD 2018 network and displayed in Figure 1.

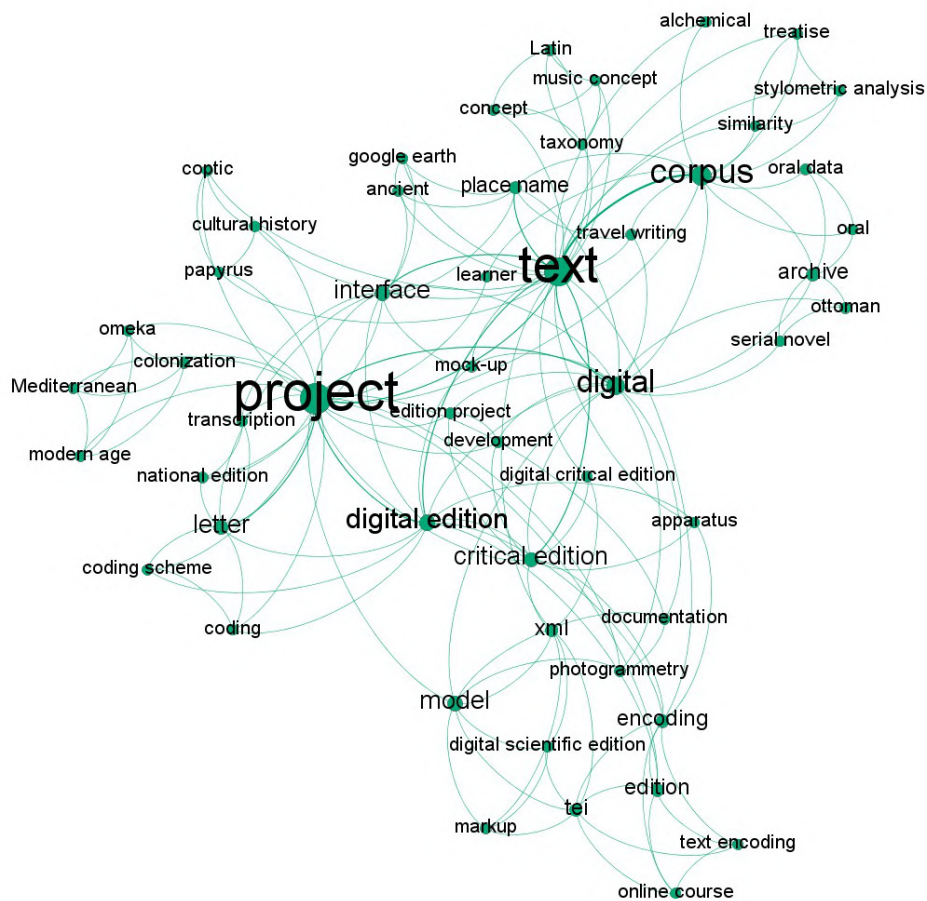


Figure 1. Biggest cluster of co-occurring keyphrases in the AIUCD 2018 network.

## BIBLIOGRAFIA

- [1] Callon, Michel, Arie Rip, and John Law, eds. Mapping the dynamics of science and technology: Sociology of science in the real world. Springer, 1986.
- [2] Kostoff, R.N., 1993. Co-word analysis. In Evaluating R&D impacts: Methods and practice (pp. 63-78). Springer, Boston, MA.
- [3] Jang, Hae Lan, Young Sung Lee, and Ji-Young An. "Application of social network analysis to health care sectors." Healthcare informatics research 18, no. 1 (2012): 44-56.
- [4] Callon, Michel, Arie Rip, and John Law, eds. Mapping the dynamics of science and technology: Sociology of science in the real world. Springer, 1986.
- [5] Bastian, Mathieu, Sebastien Heymann, e Mathieu Jacomy. «Gephi: An Open Source Software for Exploring and Manipulating Networks». In Third International AAAI Conference on Weblogs and Social Media, 361–62, 2009.
- [6] Hall, David, Dan Jurafsky, e Christopher D Manning. «Studying the History of Ideas Using Topic Models». In In Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing, Honolulu, October 2008, 363–71, 2008.
- [7] Park, Ja-Hyun, and Min Song. "A study on the research trends in library & information science in Korea using topic modeling." Journal of the Korean Society for information Management 30, no. 1 (2013): 7-32.
- [8] Ciotti, Fabio. «What's in a Topic Model? Critica teorica di un metodo computazionale per l'analisi del testo». Testo e Senso, 2017, 18.



- [9] Blei, David M. Probabilistic topic models. *Communications of the ACM*, 2012, 55.4: 77-84.
- [10] Moretti, Giovanni, Rachele Sprugnoli, and Sara Tonelli. 2015. "Digging in the dirt: Extracting keyphrases from texts with kd." In *Proceedings of CLiC-it 2015*.
- [11] Spampinato, Daria (2018) AIUCD2018 - Book of Abstracts. p. 281. ISBN 9788894253528. <http://amsacta.unibo.it/id/eprint/5997>. In: *Quaderni di Umanistica Digitale*.
- [12] Manning, Christopher, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard, and David McClosky. The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*, 2014, pp. 55-60.
- [13] Aprosio, Alessio Palmero, e Giovanni Moretti. Italy goes to Stanford: a collection of CoreNLP modules for Italian. In *arXiv preprint arXiv:1609.06204*, ArXiv prep., 2016. <http://arxiv.org/abs/1609.06204>.
- [14] Girvan, Michelle, and Mark EJ Newman. "Community structure in social and biological networks." *Proceedings of the national academy of sciences* 99, no. 12 (2002): 7821-7826.

# Hyper Spectral Imaging and the Herlufsholm Special Collection

Jakob Povl Holck<sup>1</sup>, Mogens Kragtig Jensen<sup>2</sup>, Kamilla Jensen Husen<sup>3</sup>, Anne Helle Jespersen<sup>4</sup>

<sup>1</sup> University Library of Southern Denmark, Denmark, holck@bib.sdu.dk

<sup>2</sup> University Library of Southern Denmark, Denmark, mogens.k@bib.sdu.dk

<sup>3</sup> University Library of Southern Denmark, Denmark, kamillajh@sdu.dk

<sup>4</sup> University Library of Southern Denmark, Denmark, ahj@bib.sdu.dk

In 2017, the University Library of Southern Denmark was contacted by the Danish company Newtec Engineering in Odense who wanted to test their Hyper Spectral Imaging (HSI) technology on a selection of old books containing unreadable texts in their bindings. The company develops weighing, packaging and optical sorting machinery for the food industry.

The hypothesis was that Newtec's HSI and optical scans in the near infrared spectrum would make it easier to read texts on worn medieval parchments and underneath layers of paper. Across the world, Multi Spectral Imaging and HSI have already proven themselves useful for this purpose. Therefore, a selection of monographies from the Herlufsholm Special Collection was brought to the company for optical scans, using the same kind of HSI technology that is deployed in the quality control of fruit and vegetables etc. In the experimental HSI setup, the book travels on a conveyor belt underneath special cameras and lighting. The images are sent to a pc. Among the scanned books, we have chosen to present the case of a 1583 copy of *Commentariorum de Religione Christiana Libri quatuor* by Petrus Ramus which has turned out to contain a medieval liturgical text on its cover, possibly stemming from a missal. Most likely, this fragment was fitted as part of the bookbinding by a 16<sup>th</sup> century bookbinder. Passages in the document were identified using full text searches in Google and in the Cantus Index <http://cantusindex.org/>. From a Digital Humanities perspective, the combination of HSI and Data Mining is a powerful tool when it comes to the reading and rapid identification of the above-mentioned type of fragments. It also stands to reason that both industrial companies and university libraries may benefit from these kinds of collaborations: The library is provided with an identification and the company gains technological insights.

# La *gamification* nella storia dell'arte e nella museologia: impatto sulla metodologia-

Ana Knežević

Museum of African Art, Belgrade, RS, [a.knezevic@mau.rs](mailto:a.knezevic@mau.rs)

Lo scopo di questa ricerca è quello di rappresentare e analizzare *la gamification* – *ludicizzazione* nella storia dell'arte e nella museologia, e il suo impatto sulla metodologia di queste discipline. Il termine *gamification* è apparso per la prima volta intorno al 2008, ma ha visto un'adozione diffusa solo intorno alla seconda metà del 2010 (Deterding, 2011). Altri termini che sono stati usati per descrivere questa pratica includono *giochi di produttività* (en. productivity games), *design ludico* (en. playful design), *giochi comportamentali* (en. behavioural games) e *gioco applicato* (en. applied gaming). Il termine stesso può essere definito come “l'uso di elementi di gioco e tecniche di progettazione di giochi in contesti non di gioco” (Werbach, 2012). Fra tutti gli altri contesti, la *gamification* è anche legata alle discipline di storia dell'arte e di museologia. Durante la storia recente, gli storici dell'arte e i museologi stavano usando questa tecnica per migliorare i modi di apprendere, di insegnare, di osservare e persino di collezionare, curare e interpretare gli oggetti d'arte.

Parlando di storia dell'arte, l'implementazione della *gamification* appartiene alla cosiddetta storia dell'arte digitale e la sua metodologia. In questa ricerca la differenza tra storia dell'arte *digitale e digitalizzata* viene trattata con rispetto, secondo l'opinione di Johanna Drucker – la prima implica l'uso di tecniche analitiche delle tecnologie informatiche nella storia dell'arte, mentre l'altra riguarda solo l'uso di archivi online e di basi delle immagini/ dei testi o dati rilevanti per la storia dell'arte (Zweig, 2015). In questo senso, il percorso della cosiddetta storia dell'arte digitalizzata è stato a lungo lanciato (nei primi anni settanta e ottanta, grazie a musei e biblioteche) in modo da consentire ai ricercatori un accesso più rapido e più facile al materiale che studiano. Dall'altro lato, la storia dell'arte digitale sta iniziando il suo lavoro utilizzando una metodologia informatica, per esempio - la costruzione di metadati, le analisi di rete, le analisi del discorso, la modellazione virtuale, la simulazione e la raccolta di materiali da diverse posizioni geografiche (Zweig, 2015).

In questo articolo ho intenzione di esplorare ed esaminare le caratteristiche della metodologia utilizzata in due giochi legati alla storia dell'arte – *ARTigo* e *ARTstor Arcades*. *ARTigo* è una collaborazione tra l'Istituto per l'informatica (Institute for Informatics), l'Istituto per la Storia dell'Arte (Institute for Art History) e IT-Group Humanities, tutti e tre dell'Università Ludwig-Maximilian di Monaco (<http://www.artigo.org/about.html>, 10.10.2018). *ARTigo* è un gioco realizzato secondo l'idea che gli oggetti d'arte abbiano necessità di *tagging* per facilitare gli studi di storia dell'arte e per creare automaticamente un motore di ricerca di opere d'arte (accessibile dalla piattaforma *ARTigo*). La piattaforma di gioco *ARTigo*, dall'anno 2008 offre diversi giochi che presentano una stessa opera d'arte per due giocatori (allo stesso momento). Descrivendo l'arte, lo stile, la qualità, le emozioni, questi giocatori ottengono punti quando usano lo stesso *tag*, e poi, ogni *tag* usato almeno due volte diventa una forma specifica la quale ha un ruolo importante nella classificazione delle opere d'arte su questa piattaforma (<http://www.artigo.org/about.html>, 10.10.2018).

*ARTstor Arcades* è un progetto che comprende una raccolta di quasi 125.000 fotografie di arte contemporanea non catalogate ed esposte a New York dagli anni '70 fino al 2013, quando *ARTstor* ha acquisito l'archivio (<https://artstor.blog/2015/08/31/introducing-artstors-new-crowdsourcing-software/>, 11.10.2018). *Arcades* applica il concetto di gioco al *crowdsourcing*. Offre una semplice piattaforma di gioco che consente agli utenti di inserire termini per una selezione di campi chiave per aiutare a identificare le singole opere, tra cui il titolo, l'artista, la data e la galleria. Gli utenti ricevono punti per ciascun campo di dati immesso e possono

salire progressivamente attraverso titoli prestigiosi, che vanno da “flâneur” a “master” (<https://artstor.blog/2015/08/31/introducing-artstors-new-crowdsourcing-software/>, 11.10.2018). Un giocatore riceve più punti se il nome di un artista, il titolo o un altro termine corrisponde a una risposta precedente. Questa corrispondenza è la chiave per i dati di *crowdsourcing*: più utenti, più corrispondenze, migliori saranno i dati (<https://artstor.blog/2015/08/31/introducing-artstors-new-crowdsourcing-software/>, 11.10.2018). Utilizzando una combinazione di dati e analisi degli esperti, *ARTstor* pubblicherà periodicamente immagini sufficientemente catalogate nella Biblioteca digitale *ARTstor*, dove saranno disponibili per l’insegnamento e la ricerca di oltre 500.000 utenti didattici in 1.700 college, università, scuole e musei in tutto il mondo.

Riguardo a queste descrizioni di progetti *ARTigo* ed *ARTstor Arcades*, la forma di gioco – *game* – diventa chiara; supponiamo che lo scopo sia divertente, educativo e anche utile per l’ampia comunità legata alla storia dell’arte. In altre parole, giocando con *ARTigo* ed *ARTstor Arcades*, la persona viene coinvolta nel cosiddetto *edutainment* (en. educate and entertain) – educare e divertire (Ivarsson, 2009). Ma, si può dire anche che questi giochi fanno sparire il lavoro noioso della digitalizzazione. Quando si parla della metodologia di questi giochi, le somiglianze possono essere trovate con il famoso metodo iconografico di Aby Warburg ed iconologia di Erwin Panofsky, ma solo nei suoi primi passi, cioè, gli oggetti d’arte vengono qui riconosciuti come la scena, il motivo, il soggetto o anche come l’emozione rappresentata (Dragojević, 1997). L’interpretazione finale dell’oggetto dell’arte manca – lo scopo corrente è la creazione di un grande database. Per di più, la conoscenza del soggetto della storia dell’arte rimane factografica – senza il contesto più ampio, mentre la creazione di questa abilità di *tagging* cade in grande soggettività e rischio di incompetenza. Pensando alla *gamification* nella storia dell’arte si può concludere che l’impatto dei giochi sulla metodologia di questa scienza non è quello di rivoluzionare – si tratta però di uno strumento divertente ed utile per l’apprendimento, la ricerca e la collezione delle riproduzioni e dati rilevanti per l’arte.

Da un diverso punto di vista, la *gamification* nel museo, che, secondo ICOM (International Council of Museums) è “un’istituzione permanente, senza scopo di lucro, al servizio della società, e del suo sviluppo, aperta al pubblico, che effettua ricerche sulle testimonianze materiali ed immateriali dell’uomo e del suo ambiente, le acquisisce, le conserva, e le comunica e specificamente le espone per scopi di studio, istruzione e diletto” (<http://www.icom-italia.org/definizione-di-museo-di-icom/>, 12.10.2018) può avere un ruolo più efficace. Per esempio, il gioco famoso *Ghosts of Chance* era un gioco di realtà alternativa (ARG) ospitato da Smithsonian American Art Museum nel 2008 ([http://www.museumtwo.com/publications/Ghosts\\_Report.pdf](http://www.museumtwo.com/publications/Ghosts_Report.pdf), 12.10.2018). Questo era il primo gioco di realtà alternativa al mondo a essere ospitato da un museo. Il gioco offriva al pubblico di un museo nuovo e anche esistente un nuovo modo di interagire con la collezione nella sua Luce Centro di Fondazione per l’arte americana – una struttura di stoccaggio visibile che mostra oltre 3.300 opere d’arte in teche di vetro dal pavimento al soffitto ([http://www.museumtwo.com/publications/Ghosts\\_Report.pdf](http://www.museumtwo.com/publications/Ghosts_Report.pdf), 12.10.2018). Questo tipo di giochi (ARG) sono esperienze di gioco coinvolgenti che sfumano deliberatamente la linea tra il gioco e il mondo reale. I giocatori investigano il mondo del gioco utilizzando gli stessi strumenti con cui interagiscono con il mondo reale come siti Web, e-mail, conversazioni telefoniche e anche discussioni di persona con attori che impersonificano caratteri ([http://www.museumtwo.com/publications/Ghosts\\_Report.pdf](http://www.museumtwo.com/publications/Ghosts_Report.pdf), 12.10.2018). *Ghosts of Chance* ha invitato i giocatori a creare oggetti e spedirli al museo per una mostra a cura di due personaggi del gioco in posa come dipendenti. Ma il *gioco all’interno del gioco* era anche una sfida per scoprire gli indizi della narrativa che lega quegli oggetti e investiga il modo in cui gli oggetti incarnano storie. Oltre 6.000 giocatori hanno partecipato online e 244 sono le persone

venute per l'evento sul posto ([http://www.museumtwo.com/publications/Ghosts\\_Report.pdf](http://www.museumtwo.com/publications/Ghosts_Report.pdf), 12.10.2018).

Ci sono gli altri esempi della *gamification* nella museologia – progetti come *Curator*, in cui i visitatori creano la propria collezione d'arte usando QR codici, *Framed Introspection*, a 3D portale in opera d'arte dove il visitatore può camminare ed esplorare l'opera d'arte, *Getty Thieves* – a *Twitter* gioco di realtà alternativa in cui i giocatori hanno i ruoli di ladri – amanti dell'arte e di compratori sul mercato nero, postando descrizioni e foto su *Twitter* per comprare e vendere le opere d'arte, ecc (<http://blogs.getty.edu/iris/gettyjam-report-12-museum-games-created-in-30-hours/>, 13.10.2018).

Diversamente dalla *gamification* nella storia dell'arte, questo fenomeno contemporaneo nei musei e anche nella museologia modifica i modi con cui si presenta l'oggetto museale. Il rischio di incompetenze è minore, perchè non sono tutti gli attori, ma solo coloro che lavorano nel museo – curatori, conservatori, museologi, ecc. Inoltre, il concetto curatoriale può essere sviluppato usando tecniche di *gaming* – la rappresentazione e l'interpretazione dell'oggetto del museo possono ottenere una nuova dimensione, ugualmente educativa e divertente. Nel 1938, Huizinga ha scritto: “Lo spirito della competizione ludica è, come impulso sociale, più antico della cultura stessa e pervade tutto come un vero fermento” (Huizinga, 1949). Si deve essere d'accordo con questo grande teoretico culturale, ma si deve anche riesaminare il contesto e il senso della *gamification* come un processo che porta i giochi in un contesto di non gioco (en. *non-gaming context*). È sicuro che i giochi siano utili sia nella storia dell'arte che museologia, ma è molto importante fare ulteriori riflessioni su tutta la struttura e il meccanismo di queste discipline, in questo momento preparate solo in parte per l'uso dei giochi.

Dall'altra parte dello schermo, dovrebbe essere tenuto presente che un *clic* dell'utente (quello di consumatore) non è sufficiente. Accanto ai giochi, hanno bisogno di almeno altre due: un *clic nella testa* e un *clic nella società* (Popadić, 2013). Questo *tre-clic* sarebbe utile, creativo e culturale, e la sua dimensione cognitiva sarebbe la consapevolezza, la sensibilità e la coscienza (Popadić, 2013). Solo con questo *concetto di tre-clic* possiamo cercare di creare la *gamification* abbastanza utile ed efficace nelle discipline come la museologia e la storia dell'arte.

## BIBLIOGRAFIA

- [1] Bell, D., 2006. *Cybercultures: Critical Concepts in Media and Cyber Studies*, Vol. I: Mapping Cyberculture, London and New York, Routledge.
- [2] Bentkowska-Kafel, A., 2015. „Debating Digital Art History“, in: *International Journal for Digital Art History*, Harald Klinke and Liska Surkemper (ed.), Munich, Graphentis Verlag.
- [3] Deleuze, G., Guattari, F., 1987. *A Thousand Plateaus: Capitalism and Schizophrenia*, London, University of Minnesota Press.
- [4] Deterding, S, Dixon, D, Khaled, R, Nacke, L. “From Game Design Elements to Gamefulness: Defining “Gamification””. *MindTrek '11*, September 2011.
- [5] Драгојевић, П. 1997. О развоју и елементима историје уметности, НАУКА О УМЕТНОСТИ 2, Интерна семинарска публикација.
- [6] Ivarsson, E., 2009. *Definition and prospects of the virtual museum*, Uppsala Universitet.
- [7] Foster, H., 1996. „The Archive without Museums“, in: *October*, Vol. 77, MIT Press.
- [8] Greenhalgh, M., 2004. „Art History“, in: *A Companion to Digital Humanities*, Susan Schreibman, Ray Suenebsm John Unsworth (ed.), Oxford, Blackwell.
- [9] Huizinga, J. 1949. *Homo Ludens: A Study of The Play-Element in Culture*. Routledge & Kegan Paul, Great Britain.

- [10] Lindgren, S. 2007. „From Flâneur to Web Surfer: Videoblogging, Photo Sharing and Walter Benjamin @ the Web 2.0“, in: Transformations: Walter Benjamin and the Virtual No. 15, online journal
- [11] Manovich, L., 2001. The Language of New Media, Massachusetts Institute of Technology.
- [12] Manovich, L., 2015. „Data Science and Digital Art History“, in: International Journal for Digital Art History, Harald Klinke and Liska Surkemper (ed.), Munich, Graphentis Verlag.
- [13] Popadić, M., 2013. „Totalni muzej prema Tomislavu Šoli“, u: Muzeologija, nova muzeologija, nauka o baštini – tematski zbornik, knjiga I/tom 2, Angelina Milosavljević (ur.), Beograd, Centar za muzeologiju i heritologiju.
- [14] Werbach, K, Hunter, D, 2012. For The Win: How Game Thinking Can Revolutionize Your Business. Wharton Digital Press, Philadelphia.
- [15] Zweig, B., 2015. „Forgotten Genealogies: Brief Reflections on the History of Digital Art History“, in: International Journal for Digital Art History, Harald Klinke and Liska Surkemper (ed.), Munich, Graphentis Verlag.

#### **INTERNET LINKS**

- <http://blogs.getty.edu/iris/gettyjam-report-12-museum-games-created-in-30-hours/>, 13.10.2018.
- [http://www.museumtwo.com/publications/Ghosts\\_Report.pdf](http://www.museumtwo.com/publications/Ghosts_Report.pdf), 12.10.2018.
- <http://www.icom-italia.org/definizione-di-museo-di-icom/>, 12.10.2018.
- <https://artstor.blog/2015/08/31/introducing-artstors-new-crowdsourcing-software/>, 11.10.2018.
- <http://www.artigo.org/about.html>, 10.10.2018.

# LiLa: Linking Latin. Building a Knowledge Base of Linguistic Resources for Latin

Marco Passarotti, Flavio M. Cecchini, Greta Franzini, Eleonora Litta, Francesco Mambrini, Paolo Ruffolo

CIRCSE, Università Cattolica del Sacro Cuore, Italia

## ABSTRACT

The *LiLa: Linking Latin* project was recently awarded funding from the European Research Council (ERC) to build a knowledge base of linguistic resources for Latin. LiLa responds to the growing need in the fields of Computational Linguistics and Humanities Computing to create an interoperable ecosystem of NLP tools and resources for the automatic processing of Latin. To this end, LiLa makes use of Linked Open Data (LOD) practices and standards to connect words to distributed textual and lexical resources via unique identifiers. In so doing, it builds rich knowledge graphs, which can be used for research and teaching purposes alike.

## KEYWORDS

Latin, Linguistics, Linked Open Data, Natural Language Processing, Metadata.

## INTRODUCTION

The spread of information technology has led to a substantial growth in the quantity, diversity and complexity of linguistic data available on the web. Today, although a large variety of linguistic resources (such as corpora, lexica and ontologies) exist for a number of languages, in many cases these are not interoperable. Linking linguistic resources to one another would maximise their contribution to, and use in, linguistic analysis at multiple levels, be those lexical, morphological, syntactic, semantic or pragmatic. In an ideal virtuous cycle, linguistic resources benefit from the Natural Language Processing (NLP) tools used to build them, and NLP tools benefit from the linguistic resources that provide the empirical evidence upon which they are built.

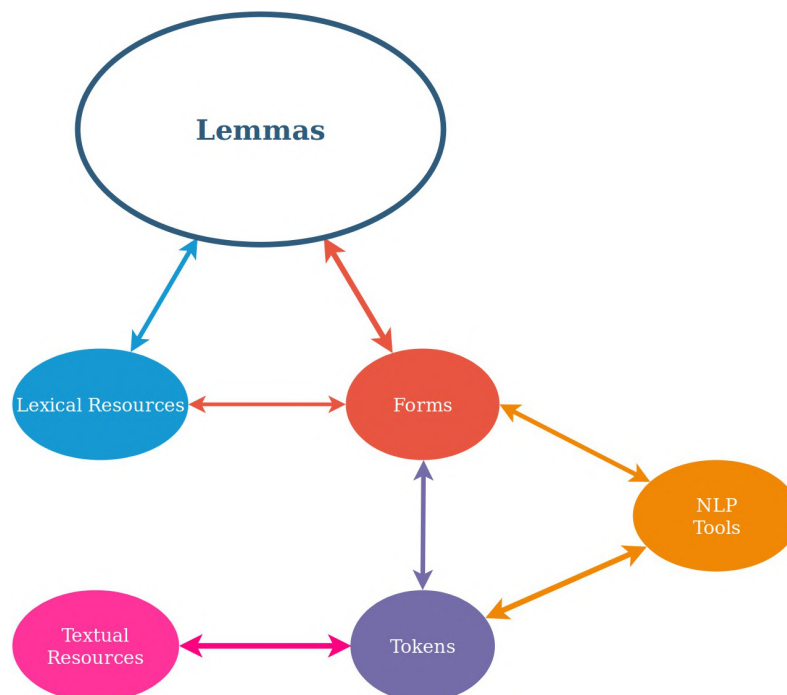
## THE PROJECT

The objective of the *LiLa: Linking Latin* project (2018-2023) is to connect and, ultimately, exploit the wealth of linguistic resources and NLP tools for Latin developed thus far, in order to bridge the gap between raw language data, NLP and knowledge description.

LiLa, which has received funding from the European Research Council (ERC) European Union's Horizon 2020 Research and Innovation programme, builds a Linked Data Knowledge Base of linguistic resources and NLP tools for Latin. The Knowledge Base consists of different kinds of objects connected via an explicitly-declared vocabulary for knowledge description. LiLa collects and connects both existing and newly-generated (meta)data. The former are mostly linguistic resources (corpora, lexica, ontologies, dictionaries, thesauri) and NLP tools (tokenisers, lemmatisers, PoS-taggers, morphological analysers and dependency parsers) for Latin. These are currently available from different providers under different licences. With regard to newly-generated (meta)data, LiLa assesses a set of selected linguistic resources by expanding their lexical and/or textual coverage. In particular, it (a) enhances the Latin texts made available by existing digital libraries and resources with PoS-tagging and lemmatisation, (b) harmonises the annotation of the three Universal Dependencies treebanks

for Latin,<sup>1</sup> (c) improves the lexical coverage of the *Latin WordNet*<sup>2</sup> and the valency lexicon *Latin-Vallex*<sup>3</sup>, and (d) expands the textual coverage of the *Index Thomisticus* Treebank<sup>4</sup>. Furthermore, LiLa builds a set of newly-trained models for PoS-tagging and lemmatisation, and works on developing and testing the best performing NLP pipeline for such a task. Connections between the aforementioned types of objects are edges labelled with a restricted set of values taken from a vocabulary of knowledge description. Connections in the Knowledge Base thus consist of a set of connections between target and source nodes.

As can be observed from the simplified conceptual model illustrated in Figure 1, the LiLa Knowledge Base is highly lexically-based. **Lemmas** are the key node type in the Knowledge Base. Lemmas occur in **Lexical Resources** (as lexical entries), but may have one or more (inflected) **Forms**. For instance, the Latin lemma *puella*, ‘girl’ has forms like *puellam*, *puellis* and *puellas*. Forms, too, can occur in lexical resources; for instance, in a lexicon containing all the word forms of a language (e.g., *Thesaurus Formarum Totius Latinitatis*, Tombeur 1998). Both Lemmas and Forms can have one or more graphical variants (*condicio* vs. *conditio*). The occurrences of Forms in real texts are **Tokens**, which are provided by **Textual Resources**. Texts in Textual Resources can be different editions/versions of the same Work (e.g., the various editions of the *Orator* by Cicero, possibly provided by different Textual Resources). Finally, **NLP tools** process either **Forms** (e.g., a morphological analyser) or **Tokens** (e.g., a PoS-tagger).



**Figure 1. Conceptual model of LiLa.**

LiLa is articulated around the following five Work Packages (WPs):

1. WP1: *Selecting and Improving Linguistic Resources for Latin*. This WP aims to provide the Knowledge Base with its backbone, i.e., linguistic data-sets.
2. WP2: *Building the Knowledge Base*. This WP represents the core of the project, as it aims to model the linguistic resources for Latin collected and selected in WP1 to build

<sup>1</sup> <http://universaldependencies.org/>

<sup>2</sup> [http://www.cyllenius.net/labium/index.php?option=com\\_content&task=view&id=21&Itemid=49](http://www.cyllenius.net/labium/index.php?option=com_content&task=view&id=21&Itemid=49)

<sup>3</sup> <https://itreebank.marginalia.it/view/lvl.php>

<sup>4</sup> <https://itreebank.marginalia.it/view/itb.php>



the Knowledge Base. Furthermore, this WP aims to make NLP tools for Latin interoperable and to connect them with linguistic resources in order to exploit the empirical evidence these provide for different NLP purposes.

3. WP3: *Querying the Knowledge Base*. This WP intends to build a user-friendly interface to allow users to write and run SPARQL queries on interconnected linguistic resources.
4. WP4: *Testing and Evaluating the Knowledge Base*. This WP tests the Knowledge Base by conducting research on its (meta)data.
5. WP5: *Disseminating the Results*. This WP is devoted to the dissemination of the results of the project through publications, conference presentations, tutorials and workshops.

This poster contribution will present the detailed structure of LiLa and describe how it meets the so-called *FAIR Guiding Principles* for scientific data management and stewardship, which state that scholarly data must be *Findable, Accessible, Interoperable* and *Reusable* (Wilkinson, 2016). In addition, it will elaborate on the progress made in WP1, which includes the evaluation of the project's PoS-tagging and lemmatisation efforts and an assessment of the resources eligible to enter the Knowledge Base, on the ontologies and vocabularies used, on the adopted URI scheme to uniquely identify nodes and their relations, and on LiLa's connection to other resources (e.g., the Linguistic Linked Open Data (LLOD) cloud).

## ACKNOWLEDGEMENTS

The LiLa project has received funding from the European Research Council (ERC) European Union's Horizon 2020 Research and Innovation programme under grant agreement No. 769994.

## REFERENCES

- [1] Tombeur, P. (1998) *Thesaurus formarum totius Latinitatis: a Plauto usque ad saeculum XXum*. Brepols, Turnhout, Belgium.
- [2] Wilkinson, M. D. et al. (2016) 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data*, 3.  
<http://dx.doi.org/10.1038/sdata.2016.18>

# Misurazione degli effetti organizzativi dei progetti digitali nelle reti documentali complesse

Brizio Tommasi

Tor Vergata University, IT, brizio.tommasi@gmail.com

## ABSTRACT

Il progetto di archivio digitale o progetto digitale è un'iniziativa ad ampio impatto organizzativo, sia di carattere sistemico che multidisciplinare; è noto in letteratura che le principali discipline coinvolte in tale contesto siano in prevalenza quattro, si fa riferimento nello specifico alle discipline archivistiche, informatiche, organizzative e giuridiche (Pigliapoco, 2016), le cui caratteristiche in un contesto di project management presentano effetti sistemici sia a livello gestionale sia a livello organizzativo. Nell'ambito dell'analisi svolta in questo paper, il progetto di archivio digitale è delineato come una gestione correlata e coerente di progetti orientati alla digitalizzazione dei flussi organizzativi documentali e alla gestione degli archivi digitali, ovvero un programma di progetti coordinati secondo i principi del multi-project management. Il progetto digitale coinvolge i diversi livelli dell'organizzazione facendo emergere una dinamica comportamentale all'interno delle reti documentali che connettono tra loro le unità di struttura dell'organizzazione, in base all'analisi della tipologia di flussi documentali che scaturiscono come effetti organizzativi del progetto digitale stesso. Il multi-project management nei contesti organizzativi sempre più digitalizzati sta rilevando una trasformazione degli aspetti organizzativi e decisionali, facendo emergere nuove caratteristiche sia nei processi di propagazione informativa sia nei processi di collaborazione e di tipo decisionale, che possono essere rappresentate e analizzate attraverso lo studio della complessità che li caratterizza. Si fa riferimento in particolare alla cosiddetta organizational network analysis (Anderson, 1999); si tratta nello specifico dello studio di reti complesse (Barabási, 2002) emergenti dall'analisi dei componenti della rete documentale identificati in questo paper in forma aggregata o ricorsiva.

## PAROLE CHIAVE

digital archive, project management, records management, digital network, organization

## INTRODUZIONE

Le reti documentali complesse consentono di rappresentare e analizzare le interazioni tra le unità organizzative che caratterizzano la gestione degli archivi digitali dei documenti, come parte integrante di uno degli elementi comportamentali della *digital humanities*. Le proprietà specifiche dei singoli elementi non sono considerate (almeno in prima battuta) in favore di un'attenzione sulla struttura globale delle interazioni documentali che operano nell'archivio digitale; ciò permette di rilevare aspetti e comportamenti complessivi che altrimenti rimarrebbero nascosti, consentendo dunque di riconoscere macrostrutture o i cosiddetti *pattern* che includono elementi e comportamenti della rete in prima analisi apparentemente non correlati. Una rappresentazione della rete di relazioni documentale è riportata nella seguente figura in forma di grafo orientato i cui nodi sono le unità organizzative della rete e gli archi le relazioni documentali della rete; le differenti dimensioni e i differenti colori dei nodi caratterizzano l'analisi quantitativa dei flussi e delle relazioni documentali della rete complessa.

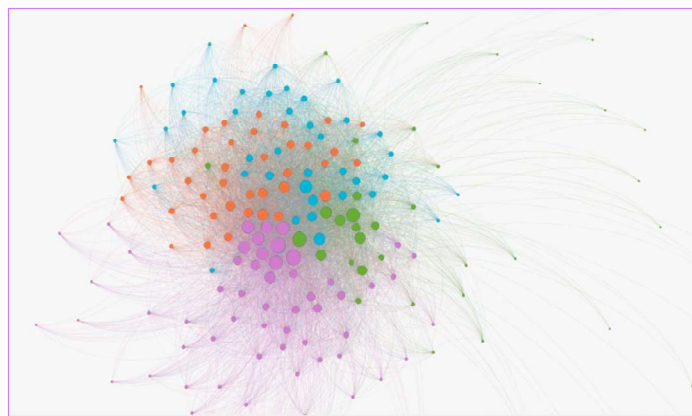


Figura 1 – Una rappresentazione della rete documentale complessa.

Le reti di relazioni documentali sono analizzate, in questo paper, considerando le componenti organizzative della rete in forma aggregata, ovvero si considerano sia i flussi documentali (archi) suddivisi e aggregati in base alle tipologie di flusso, sia le unità organizzative (nodi) suddivise e aggregate in base al livello operativo nell'ambito dell'organizzazione. Nello specifico sono analizzate due classi di reti documentali con soggetti aggregati: (i) La rete dei "cicli di lavoro documentale" in cui si considerano nell'analisi le seguenti tre tipologia di flusso documentale (archi): assegnazione visibilità; assegnazione in copia; assegnazione con notifica; che caratterizzano la propagazione informativa documentaria all'interno dell'organizzazione. (ii) La rete dei "flussi di collaborazione organizzativa" in cui si considerano nell'analisi le seguenti sei tipologia di livello operativo delle unità di struttura dell'organizzazione: Direzione; UO di livello1; Segreteria; Staff; UO di livello2; Utenti. Passiamo ora ad analizzare singolarmente le due reti di interconnessione emergenti dall'analisi dei flussi aggregati e delle unità organizzative aggregate, rappresentando i risultati scaturiti dalle differenti tipologie di dati.

## CICLI DI LAVORO DOCUMENTALE NEGLI ARCHIVI DIGITALI

Nell'analisi dei cosiddetti cicli di lavoro dei flussi documentali, che definiamo life-cycle network, si considerano i flussi operativi in forma aggregata rispetto alle relazioni documentarie (archi). Ciò al fine di analizzare e identificare l'effetto della propagazione informativa dei flussi documentali in funzione della numerosità delle differenti tipologie di flusso prese in esame. La rete dei cicli di lavoro rappresenta, dunque, il ciclo di vita dei documenti nell'ambito della loro trattazione amministrativa, in termini di propagazione dei flussi documentali all'interno dell'organizzazione. In base a questo è possibile calcolare un possibile "indice di propagazione" dei flussi documentali, che emerge sin in modo "esplicito" che in modo "implicito". In altri termini se il ciclo di flusso è notificato si considera come grado di propagazione esplicito, altrimenti se il ciclo di flusso aggregato è gestito come visibilità o come copia conoscenza si considera come grado di propagazione implicito. Si osserva come l'indice di propagazione oscilla mediamente tra circa 2-3 (con notifica o esplicito) e 5-6 (senza notifica o implicito), dove si rileva nel grado esplicito le attività di propagazione tipiche dei flussi documentali archivistici: smistamento, assegnazione, sotto-assegnazione; mentre nel grado implicito la distanza massima dei percorsi di propagazione informativa documentaria; equivalente al massimo grado di propagazione (Milgram, 1967) che caratterizza i cicli di flusso documentale.

Nella tabella seguente sono riportate le rilevazioni effettuate per la misurazione dei cicli di lavoro con notifica e di quelli totali, evidenziati nelle due colonne (a sfondo grigio); mentre in

ogni colonna sono riportati i dati rilevati (prime sei colonne) e le formule applicate sui dati (dalla colonna 7 alla 14).

1	2	3	4	5	6	7	8	9	10	11	12	13	14
-	-	-	Flussi automatici	Flussi gestionali	Documenti	% Doc / Assegnaz. =>	Con notifica	Senza notifica	In Copia	sub-totali	con notifica + senza + copia	Workflow / Doc	% (Doc+Wf) / Assegnazione
Tipologia Assegnazione (x)	Anno Y	Totali A	Workflow W	Assegnazioni A(x)	Tot. Doc. D	Incidenza del ciclo di lavoro $I_1 = 1 - (D/A_n) \%$	Cicli di lavoro $F = A(n)/D$	Cicli collaborativi $C1 = A(v)/D$	Cicli conoscitivi $C2 = A(c)/D$	Cicli estesi $E = C1 + C2$	Cicli di lavoro TOTALI $T = F + E$	% Flussi automatici $W_f = W/D$	Incidenza ciclo TOTALE $I_2 = 1 - (D + W)/A$
Visibilità (v)	½ 2013	16528	1	16527									
Copia (c)	½ 2013	57915	4732	53183									
Notifica (n)	½ 2013	141289	4768	136521	44049	67,7%	3,1	0,4	1,2	1,6	4,7	11%	64,2%
Visibilità (v)	2014	219004	17	218987									
Copia (c)	2014	135788	13924	121864									
Notifica (n)	2014	290531	13927	276604	101835	63,2%	2,7	2,2	1,2	3,3	6,1	14%	58,1%
Visibilità (v)	2015	64636	471	64165									
Copia (c)	2015	139066	9129	129937									
Notifica (n)	2015	279300	9256	270044	152770	43,4%	1,8	0,4	0,9	1,3	3,0	6%	40,0%
Visibilità (v)	½ 2016	152554	3	152551									
Copia (c)	½ 2016	125721	11870	113851									
Notifica (n)	½ 2016	212553	11873	200680	84779	57,8%	2,4	1,8	1,3	3,1	5,5	14%	51,8%

Tabella 1 – Propagazione documentale e incidenza dei relativi cicli di lavoro.

Con riferimento ai risultati riportati nella tabella precedente, si riportano qui di seguito le formulazioni adottate e la relativa descrizione:

Dato	Descrizione	Formulazione
Tipologia assegnazione	Assegnazione dei documenti secondo tre modalità (x): visibilità, copia, notifica.	(v) Visibilità documenti (c) Visibilità in Copia (n) Visibilità con Notifica
Anno	Anno di osservazione dei dati del progetto di archivio digitale: ½ 2013, 2014, 2015, ½ 2016	Y
Totali	Totali delle assegnazioni documentali suddivise per tipologia	A
Workflow	Flussi automatici di assegnazione dei documenti gestiti dal software	W
Assegnazioni	Flussi gestionali di assegnazioni dei documenti gestiti dagli utenti	A(x)
Tot. Doc.	Totale documenti prodotti nel periodo, indipendente dal tipo di assegnazione	D
Incidenza del ciclo di lavoro	Percentuali assegnazioni gestite con notifica dagli utenti al netto dei documenti prodotti	$I_1 = 1 - (D/A_{(n)})$
Cicli di lavoro	Numero di assegnazioni con notifica rispetto al numero dei documenti prodotti	$F = A_{(n)}/D$
Cicli collaborativi	Numero di assegnazioni senza notifica rispetto al numero dei documenti prodotti	$C1 = A_{(v)}/D$
Cicli conoscitivi	Numero di assegnazioni in copia conoscenza rispetto al numero dei documenti prodotti	$C2 = A_{(c)}/D$
Cicli estesi	Totale delle assegnazioni senza notifica e in copia conoscenza	$E = C1 + C2$
Cicli di lavoro Totali	Totale complessivo assegnazioni con notifica, senza notifica e in copia conoscenza	$T = F + E = F + C1 + C2$
% Flussi automatici	Percentuale dei flussi di assegnazione workflow	$W_f = W/D$

	rispetto al totale dei documenti prodotti	
Incidenza ciclo Totale	Percentuale delle assegnazioni con notifica totali al netto dei documenti prodotti	$I_2 = I - (D+W)/A_{(n)}$

Tabella 2 – Descrizione dei dati relativi ai flussi di propagazione documentale e i relativi cicli di lavoro.

Attraverso tali rilevazioni è possibile ipotizzare alcuni parametri di misurazione e di controllo del grado di efficacia e di efficienza nel trattamento gestionale dei documenti in base alla loro propagazione. Se il grado di propagazione esplicito supera la soglia critica 3, può esserci un elemento di inefficacia dei criteri di assegnazione documentale o nella corrispondente struttura organizzativa. Se il grado di propagazione implicito supera la soglia critica 6, si possono evidenziare elementi organizzativi propensi ad un “eccesso” di propagazione informativa, facendo rilevare ad esempio fattori di rischio nell’ambito di eventuali criteri di riservatezza documentaria; si rilevano anche elementi di eccessiva attività operativa nell’organizzazione nell’ambito ad esempio dei processi interni o dei procedimenti amministrativi.

## FLUSSI DI COLLABORAZIONE ORGANIZZATIVA NEGLI ARCHIVI DIGITALI

La misurazione della cosiddetta “omofilia” nell’analisi delle reti consente di effettuare valutazioni previsionali al fine di identificare quelle componenti della rete organizzativa che tendono a comportarsi in maniera uniforme in quanto identificate da caratteristiche uniformi (McPherson et al., 2003). In particolare per l’analisi delle relazioni di collaborazione organizzativa, nella propagazione dei flussi documentali, si sono presi in considerazione i nodi della rete organizzativa in forma aggregata (strutture organizzative funzionali o formali) rispetto alle relazioni documentarie. Di seguito si rappresentano i criteri di calcolo dell’indice di collaborazione tra unità organizzative aggregate, sia come mittente (*source*) che come destinataria (*target*), sulla base del rapporto tra la sommatoria dei flussi documentali complessivi rispetto al totale delle relazioni documentali tra i nodi aggregati della rete, seguendo la seguente formulazione:

$$\text{Indice di collaborazione organizzativa} = \frac{\sum \text{flussi documentali complessivi}}{\sum \text{relazioni documentarie}}$$

Il calcolo di questo rapporto tra il totale dei flussi documentali scambiati e il totale delle relazioni organizzative documentali, ha consentito di ipotizzare una possibile misurazione del grado di “omofilia organizzativo causale”, portando ad identificare quelle unità organizzative aggregate che hanno una forte relazione, anche causale rispetto alle regole di propagazione informativa aziendale.

Nello specifico, in tabella sono rappresentati gli indici di collaborazione organizzativa, per ogni livello funzionale della struttura analizzata, da cui è possibile ipotizzare i fattori di omofilia organizzativa nei valori di maggiore evidenza.

		Target								
		⇒								
		Direzion e	UO liv.1	Segreteri a	Staff	UO liv.2	Utenti	Σ	Max	
Source	<u>Omofilia</u>									
↓	Direzione	0,0	7,1	<u>7,5</u>	2,0	3,5	3,3	23,4	7,5	
	UO liv.1	37,6	5,9	<u>229,8</u>	18,6	40,8	97,1	429,9	8	

<b>Segreteri a</b>	151,1	<b>308,3</b>	204,8	238, 8	316,1	<b>395,6</b>	1614, 7	<b>395, 6</b>
<b>Staff</b>	<u>134,3</u>	17,1	113,7	6,7	28,1	20,4	<b>320,2</b>	<b>134, 3</b>
<b>UO liv.2</b>	22,1	141,5	150,4	63,4	16,9	<u>210,2</u>	<b>604,4</b>	<b>210, 2</b>
<b>Utenti</b>	23,4	77,1	<b>161,8</b>	27,6	90,5	38,8	<b>419,1</b>	<b>161, 8</b>
$\Sigma$	<b>368,4</b>	<b>557,1</b>	<b>867,9</b>	<b>357, 1</b>	<b>495,9</b>	<b>765,4</b>		
<b>Max</b>	<b>151,1</b>	<b>308,3</b>	<b>229,8</b>	<b>238, 8</b>	<b>316,1</b>	<b>395,6</b>		

Tabella 3 – Indice di collaborazione organizzativa e fattori di omofilia causale.

Si rileva in particolare l'emergere di due principali fattori di omofilia nel comportamento organizzativo (Axelrod, 1984):

1. Lo smistamento e l'assegnazione documentale tra UO di livello 1, gerarchicamente superiore, e la Segreteria in entrambe le direzioni dei flussi documentali;
2. L'assegnazione e la consegna, o sotto-assegnazione, documentaria tra Segreteria e Utenti in entrambe le direzioni dei flussi documentali.

Le rilevazioni di cui al precedente calcolo dell'indice di collaborazione e di omofilia organizzativa, sono stati rapportati, nel periodo di osservazione della rete documentale, il totale degli archi (relazioni documentarie) che interconnettono tra le unità organizzative aggregate, rispetto al totale complessivo dei pesi relativi ai flussi documentali rilevati sui medesimi archi (flussi documentali complessivi), le cui evidenze sono riportate nelle tabelle seguenti.

Nella prima tabella sono rappresentate le quantità di relazioni documentarie (archi della rete) tra i nodi aggregati della rete documentale, unitamente alla sommatoria e al valore massimo rilevabile per righe e per colonne delle unità organizzative aggregate, ovvero sia come mittente (*source*) che come destinataria (*target*).

		Target ⇒								
		UO liv.1		Segreteria	Staff	UO liv.2		Utenti	$\Sigma$	Max
Source ↓	$\Sigma$ ARCHI	Direzione								
	<b>Direzione</b>	0	7	13	1	10	3	<b>34</b>	<b>13</b>	
	<b>UO liv.1</b>	19	36	92	20	106	73	<b>346</b>	<b>106</b>	
	<b>Segreteria</b>	52	136	418	125	458	509	<b>1698</b>	<b>509</b>	
	<b>Staff</b>	15	47	124	41	177	174	<b>578</b>	<b>177</b>	
	<b>UO liv.2</b>	54	166	362	112	509	530	<b>1733</b>	<b>530</b>	
	<b>Utenti</b>	78	231	658	249	960	1161	<b>3337</b>	<b>1161</b>	
	$\Sigma$	<b>218</b>	<b>623</b>	<b>1667</b>	<b>548</b>	<b>2220</b>	<b>2450</b>			
	<b>Max</b>	<b>78</b>	<b>231</b>	<b>658</b>	<b>249</b>	<b>960</b>	<b>1161</b>			

Tabella 4 – – Calcolo del numero totale di relazioni documentali (archi) tra i nodi aggregati.

Nella seconda tabella sono rappresentate le quantità di flussi documentali complessivi (pesi

degli archi) tra i nodi aggregati della rete documentale con la sommatoria e il valore massimo rilevabile per righe e colonne delle unità organizzative aggregate, sia come mittente (*source*) che come destinataria (*target*).

		Target ⇒						Σ	Max
		Direzione	UO liv.1	Segreteri a	Staff	UO liv.2	Utenti		
Source ↓	Σ PESI								
	Direzione	0	50	97	2	35	10	194	97
	UO liv.1	714	214	21146	372	4325	7089	33860	21146
	Segreteri a				2984	14479	20137	51142	20137
	Staff	7855	41931	85619	7	0	8	0	8
		2014	803	14098	275	4967	3551	25708	14098
							11140	20620	11140
	UO liv.2	1192	23489	54427	7100	8590	2	0	2
	Utenti	1826	17807	106439	6870	86881	45009	26483	10643
					4446	24958	36843		
	Σ	13601	84294	281826	6	8	9		
					2984	14479	20137		
	Max	7855	41931	106439	7	0	8		

Tabella 5 – Calcolo del numero totale di flussi documentali (pesi) tra i nodi aggregati.

## b. GUADAGNO INFORMATIVO NEI FLUSSI DOCUMENTALI

Pariamo dapprima dall'ipotesi che in una rete organizzativa complessa lo scambio di flussi documentali, emergenti dal progetto di archivio digitale nei suoi effetti organizzativi, si ottiene un "guadagno informativo" dai flussi documentali (pay-off), ottenuti grazie ai cicli di lavoro dei documenti trasmessi tra le strutture organizzative (von Neumann et al., 1944). Riprendendo i dati sull'indice di collaborazione, nella tabella finale sono rappresentati i guadagni di flussi rispettivamente per le unità mittenti Source (righe) e per le unità destinatarie Target (colonne), ottenuto mettendo in un'unica rappresentazione i dati combinati relativi ai precedenti indici di collaborazione organizzativa.

		Target ⇒					
		EQUILIBRIO Direzione	UO liv.1	Segreteria	Staff	UO liv.2	Utenti
Source ↓							
	Direzione	(0 - 0)	(7 - 38)	(7 - 151)	(2 - 134)	(4 - 22)	(3 - 23)
	UO liv.1	(38 - 7)	(6 - 6)	(230 - 308)	(19 - 17)	(41 - 142)	(97 - 77)
	Segreteria	(151 - 7)	(308 - 230)	(205 - 205)	(239 - 114)	(316 - 150)	(396 - 162)
	Staff	(134 - 2)	(17 - 19)	(114 - 239)	(7 - 7)	(28 - 63)	(20 - 28)
	UO liv.2	(22 - 4)	(142 - 41)	(150 - 316)	(63 - 28)	(17 - 17)	(210 - 91)
	Utenti	(23 - 3)	(77 - 97)	(162 - 396)	(28 - 20)	(91 - 210)	(39 - 39)

Tabella 6 – Propagazione documentale e incidenza dei relativi cicli di lavoro.

Considerando, pertanto, lo scambio di flussi documentali come “guadagno” di flusso informativo delle strutture organizzative aggregate, sembra emergere un fattore di equilibrio (Nash, 1951) proprio nei punti di relazione ad alta omofilia organizzativa come definita in precedenza.

## **CONSIDERAZIONI SUI PARAMETRI DI MISURAZIONE DELLE RETI DOCUMENTALI**

Il progetto digitale e il programma di trasformazione digitale dell'organizzazione, sia in ambito pubblico che privato, stanno delineando una forma di complessità all'interno delle organizzazioni orientate al digitale, con la conseguente identificazione di reti organizzative complesse nella gestione degli archivi digitali come elementi caratterizzati nell'ambito della *digital humanities*. Nel corso delle attività di ricerca realizzate, sono state analizzate le reti emergenti per misurare gli effetti del programma di progetti digitali multipli. In questo paper possiamo rappresentare i seguenti principali risultati: (1) L'emergere di una forma di equilibrio tra i nodi aggregati nella rete organizzativa si basa sui flussi documentali propagati all'interno dell'organizzazione. Questo equilibrio dipende anche dal grado di omofilia organizzativa, come precedentemente definita, e ne definisce un “pay-off informativo”. (2) Il grado di propagazione dei flussi documentali all'interno dell'organizzazione definisce un indice esplicito e uno implicito sull'efficacia del ciclo di vita documentale. (3) L'incidenza del ciclo di vita documentale per la gestione manuale o automatica della rete di flussi di documenti può identificare possibili fattori di efficienza nella gestione di documenti e flussi di lavoro.

I sistemi complessi hanno un'evoluzione dinamica ed è necessario misurare periodicamente le prestazioni della rete organizzativa digitale: ogni sei mesi o ogni anno. In generale queste misure sono complesse e ci sono difficoltà ad analizzarle con metodo analitico. È utile effettuare misurazioni mediante analisi di rete che ci consentono di avere una visione generale della gestione comportamentale in un programma di progetti digitali.

Possiamo dire che le reti decisionali basate sui flussi di documenti organizzativi nei progetti digitali definiscono un processo di "auto-organizzazione" (De Toni, 2011). Le comunicazioni formali nelle reti mettono in evidenza una struttura predefinita. Il complesso delle comunicazioni formali e informali mette in evidenza una complessa rete di documenti con differenti centralità e comportamenti. Questi modelli di rete possono permettere un'analisi del processo di evoluzione delle reti decisionali, insieme ai possibili fattori per la degenerazione incontrollata (Ashby, 1947) dei flussi documentali nell'organizzazione. In questo modo possiamo identificare le reti di flussi di documenti più resistenti ai continui cambiamenti nell'organizzazione. Inoltre possiamo ottimizzare le reti esistenti per un miglioramento continuo dell'efficienza organizzativa.

Dalle analisi dei risultati raggiunti, si possono considerare infine alcuni elementi utili alle finalità della ricerca; si fa riferimento in particolare ai parametri relativi alla misurazione del numero minimo e del numero massimo di assegnazioni documentali, unitamente al tasso di incidenza dei documenti rispetto alle assegnazioni, oltre alla misurazione del numero di assegnazioni tra le unità aggregate per livello gerarchico. Questi parametri possono essere adottati quali criteri per la valutazione degli impatti dei flussi documentali nell'organizzazione, al fine di utilizzarli come fattori decisionali nelle scelte dei progetti da



avviare e schedulare nell'ambito del portafoglio aziendale relativo alle iniziative di trasformazione digitale<sup>1</sup>.

## **BIBLIOGRAFIA**

- [1] Anderson P.W., 1999, "Complexity Theory and Organization Science", *Organization Science: A Journal of the Institute of Management Sciences*, Vol. 10, N. 3, pp. 216-232.
- [2] Ashby W.R., 1947, "Principles of the Self-Organizing Dynamic System", *Journal of General Psychology*, Vol. 37, pp. 125-128.
- [3] Axelrod R., 1984, "The Evolution of Cooperation", Basic Books.
- [4] Barabàsi A.L., 2002, "Linked, The New Science of Networks", Basic Books.
- [5] De Toni A., et al., 2012, "Auto-organizations", Marsilio.
- [6] McPherson M., Smith-Lovin L., Cook J. M., 2003, "Birds of a Feather: Homophily in Social Networks". *Annual Review of Sociology*, Vol. 27, pp. 415-444.
- [7] Milgram S., 1967, "The Small World Problem", *Psychology Today*. Vol. 2, pp. 60-67.
- [8] Pigliapoco S., 2016, "Progetto archivio digitale. Metodologia, sistemi, professionalità", Civita, Lucca.
- [9] von Neumann J., Morgenstern O., 1944, "Theory of Games and Economic Behavior", Princeton University Press.
- [10] Nash J. F., 1951, "Non-Cooperative Games", *The Annals of Mathematics*, Vol. 54, pp. 286-295.

---

<sup>1</sup> Le opinioni espresse dall'autore sono personali e non impegnano il proprio Istituto.

# Strolling around the dawn of Digital Humanities

Gabriella Pardelli, Sara Goggi, Federico Boschetti

CNR, Istituto di Linguistica Computazionale, "Antonio Zampolli", Italy gabriella.pardelli@ilc.cnr.it, sara.goggi@ilc.cnr.it, federico.boschetti@ilc.cnr.it

In questo contributo tentiamo di fissare alcuni momenti salienti del processo che ha visto la nascita comune della Linguistica Computazionale e delle Digital Humanities nonché i loro alterni allontanamenti e ricongiungimenti.

Computer science is a recent discipline, without a clear epistemological statute, born out of a number of open interdisciplinary fields throughout and immediately after the Second World War. Bletchley Park, where the machines for decoding the messages of the German Army, and the Macy's Lectures (1946–1953), where the idea of cybernetics was born, are cases where the transdisciplinary nature of informatics is most strongly evident. (Fiormonte et alii, 2000) [1]

Nelle ricerche umanistiche l'impiego dell'elaboratore elettronico prende il via nella seconda metà del ventesimo secolo favorendo l'uso di metodi statistici sia nello studio di opere letterarie che nello studio delle lingue, promuovendo un sodalizio interdisciplinare che è arrivato ai giorni nostri senza interruzione.

Sin da subito fu anche avvertita l'esigenza di recuperare una terminologia esaustiva per indicare l'incontro dei sistemi di automazione con le analisi linguistico-letterarie. Occorreva legare l'anima cibernetica con quella umanistica. Dalla fine degli anni '50 ad oggi è possibile intravedere nel corso dei decenni rimodulazioni nei diversi ambiti di ricerca percepibili e inaugurati da espressioni linguistiche inedite. In quegli anni il lavoro riguardante la traduzione automatica, *machine translation*, *mechanical translation* era iniziato, in special modo per lingue come l'inglese e il russo; altresì per la prima volta, l'uso degli elaboratori elettronici nello spoglio di testi viene avviato da Padre Busa nel 1951.

Per una suddivisione temporale delle attività svolte dal settore retrospettivamente ci concentriamo su alcune porzioni di testo tratte dall'universo o docuverso di contributi in libri in riviste e in atti di convegno sulle applicazioni dell'elaborazione elettronica alle scienze umane. Citiamo due autori, Kenny e la Terras, che tratteggiano in decenni la storia dell'accostamento dell'elettronica alle *Humanities*.

Kenny:

*In all humanities discipline the use of computers seems to have passed through three stages, each corresponding roughly to a decade.* (Kenny, 1992) [2].

Il primo nucleo pionieristico iniziato con Busa e la sua compilazione automatica delle concordanze delle opere di San Tommaso d'Aquino all'inizio degli anni '50 consacrò una esplosione di applicazioni nelle ricerche umanistiche che non si è più fermata.

Nel 1962 Bernard Quemada scrive:

*L'utilisation des machines a introduit dans notre ordre de recherches un bouleversement total. Des travaux de type artisanal, nous passons brutalement aux techniques industrielles pour lesquelles rendement et prix de revient entrent en première ligne de considération. La machine nous impose, dès lors, des démarches collectives soumises à une planification rigoureuse que les spécialistes devraient rapidement élaborer.* (Quemada, 1962) [4].

Nei primi anni '60 proseguono le prime conferenze del giovane settore identificato dagli addetti ai lavori con differenti termini. A tal proposito ricordiamo due colloqui europei e una conferenza americana:

1. *Besançon, 1961* [4]: *Colloque International sur la Mécanisation des Recherches Lexicologiques*. Il focus del colloquio fu quello di riunire gli studiosi delle varie discipline umanistiche per mettere in luce le possibilità offerte dal nuovo mezzo <meccanografico> attraverso il confronto delle esperienze di ricerca.

Riportiamo dalla sezione di Lessicografia e Stilistica le parole di Jean Petit riferite allo spoglio elettronico di una <pièce de théâtre> delle due versioni de *La Ville* di Claudel:

*Ainsi dans la second version de La Ville l'étude des images musicales, physiologique et lumineuses conduit à toute une interprétation du drame.* (Petit, 1961) [14], pagina 169.

Jecques Petit sottolinea l'importanza dell'indicizzazione elettronica ai fini della comparazione delle due versioni del dramma di Claudel, osservando che uno studio tradizionale non avrebbe mai permesso il recupero di elementi nuovi dell'opera teatrale.

2. *Praga (1966)* [5]; *Les Machines dans la Linguistique*. *Colloque International sur la Mécanisation et l'Automation des Recherches Linguistiques*. Pochi anni dopo *Besançon* recuperiamo termini come <macchine>, <meccanizzazione>, <automazione>.

Al colloquio vengono presentate le esperienze di ricerca dei laboratori europei del settore. In particolare l'analisi dei testi. Per l'Italia, Busa, Duro e Zampolli

Con chiarezza Busa sottolinea le due fasi preliminari delle ricerche linguistiche:

- a) La transcription de textes naturels, à partir du livre imprimé, sur bande magnétique en vue d'élaborations électroniques...
- b) b)...la compilation d'indices et de concordances de mots, de morphèmes, de graphèmes, de syntagmes, de fréquences, etc...(Busa, 1966) [15].

3. New York, 1964 [6]; *Literary Data Processing Conference*. Il titolo della conferenza costituito da una tripletta di termini risulta essenziale e efficace per indicarne il dominio disciplinare. Dall'evento recuperiamo la *querelle* tra Parrish e McDonough sul legame degli umanisti con le macchine:

*We are not asking the computer to be a critic; we are asking it to help us to be better critics.* (Parrish, 1964) [7].

*The machine no more replaces the scholar than the Renaissance patron replaced the artist: each merely provides the leisure for truly human activity. We do not wish machines to act as critics, as Professor Parrish said, only to save critics from acting as machines.* (McDonough, 1964) [36].

Significativa la riflessione introduttiva di Bessinger:

*... some humanists who have adopted computers are not yet fully adapted to them.* (Bessinger, 1964) [2].

Nel 1968 Froger nell'introduzione al libro sull'automatizzazione della critica testuale si esprime con queste parole:

*Il est maintenant courant d'automatiser les recherches philologiques ou littéraires, et d'utiliser la machine électronique ou mécanographique pour des études lexicographiques et stylistiques; ces procédés ont déjà été employés aussi pour aider la critique conjecturale, et par exemple combler les lacunes des manuscrits de la mer Morte; mais on ne les avait pas appliqués jusqu'ici aux opérations de la critique textuelle qui consistent à collationner les manuscrits et à déterminer leur relation généalogiques.* (Froger, 1968) [7].

E della nascita della rivista dedicata al settore *Computers in the Humanities* Froger entusiasticamente commenterà:

*... une revue spécialisée, Computers in the Humanities, énumère cent vingt programmes en cours d'exécution et l'énumération n'est pas exhaustive* (Froger, 1968). [7].

Eloquenti sono le considerazioni scritte nel primo numero della rivista da Milic:

*We define humanities as broadly as possible. Our interests include literature of all times and countries, music, the visual arts, folklore, the non-mathematical aspects of linguistics, and all phases of the social sciences that stress the human. When, for example, the archeologist is concerned with fine arts of the past, when the sociologist studies the non-material facets of culture, when the linguist analyzes poetry, we may define their intentions as humanistic; if they employ computers, we wish to encourage them and to learn from them... With the inauguration of Computers and the Humanities, the time has perhaps arrived for a more serious look at the position of the humanistic scholar in the world of data processing* (Milic, 1966) [12].

Nel 1967, tra gli addetti ai lavori, grande emozione suscitò la terza generazione dei computer contrassegnata dall'uso di circuiti integrati e Wishy in un articolo sull'hardware per le Humanities, con entusiasmo rilevò:

*There is no doubt that the third generation computers will eventually measure up to much that was promised. They are extraordinarily fast and can reach enormous amounts of data in immediately available storage; performance of more than one program at a time is feasible; the diversity of special devices such as voice response and graphic mechanisms that can be incorporated into a system are most impressive* (Wishy, 1967) [13].

Nel secondo volume della rivista molto spazio è dedicato al software, in aggiunta viene pubblicato una sorta di report molto dettagliato intitolato *Computer Programs Designed to Solve Humanistic Problems*. Recuperiamo dal *Prospect*:

*Among "Computer Programs," reads will find at least one for structural analysis, for various types of musicological analysis, for concordance-making, and for creating visual designs from mathematical formulae.* (*Computers and the Humanities*, 1966) [16].

Mesthene nel 1969 si interroga sul rapporto tra tecnologia e scienze umane sul piano etico e sociale:

*"What, now, is the effect of technology on that domain? How are the humanistic disciplines dealing with the spectacular elaboration and enrichment of our society's toolbox? Put the other way, with what challenges does modern technology confront our humanists, our artists, and our theologians? I want to suggest to you that the*

*challenges are significant, that they are poorly understood, and poorly met when they are met at all".* (Mesthene, 1969) [8].

Nello stesso anno una voce femminile, quella della Montgomery, si leva a difesa dell'utilità della scienza informatica:

*"Whether or not one accepts the idea of the so-called "information explosion, " the processing of natural language text is an important challenge for both linguistics and computer science... A further--and not unrelated--reason for non-participation linguists in automated language processing is a basic lack of knowledge about computers in the sense of realizing when a computer is a handy tool and when it isn't so handy"* (Montgomery, 1969) [9].

Per Kenny il periodo che segue quello euristico e pionieristico sembra aver raggiunto un equilibrio in parte somministrato da una maggiore specializzazione degli umanisti in campo informatico:

*The second period lasted through the seventies until the early eighties. Humanities scholars learned now to write their own programs, as programming languages became easier to master.* (Kenny, 1992). [2].

A seguire la terza decade, contrassegnata dall'avvento del personal computer e quindi dalla crescente, anche se non completa, indipendenza dai centri di calcolo:

*Throughout the seventies and the early eighties, humanities scholars, despite their growing independence, still worked collaboration with, and usually in proximity to, their colleagues in computing centres.* (Kenny, 1992) [2].

Nel 1980 presso la Scuola Normale Superiore di Pisa si svolse la prima conferenza internazionale relativa alle applicazioni informatiche nelle ricerche sulla storia dell'arte, archeologia, architettura, catalogazione e indicizzazione di archivi, bibliografia, rilevamento di furti d'arte e iconografia [10]

Nel 1986 nasce la rivista *Literary and Linguistic Computing, LLC* di 'ALLC, che dal 2015 diventerà *Digital Scholarship in the Humanities, DSH di ADHO e EADH*.

Terras:

Il libro scritto dalla Terras fornisce elementi di discussione sull'uso di certe tecnologie in campo umanistico a partire dal nuovo termine Digital Humanities.

*The term Digital Humanities has been commonly used to describe the application of computational methods in the arts and humanities for 10 years, since the publication, in 2004, of the Companion to Digital Humanities. "Digital Humanities" was quickly picked up by the academic community as a catch-all, big tent name for a range of activities in computing, the arts, and culture.* [Terras, 2016] [3]

A veder bene l'elaborazione elettronica rivolta allo studio delle scienze umane <Digital Humanities> (DH) dal 2004, e la <Machine dans la Linguistique> <Computational Linguistics> (CL) dal 1966<sup>1</sup> storicamente hanno un background condiviso, quello dell' *information explosion*.

Le radici comuni si intravedono ancora oggi. Se leggiamo con attenzione la citazione di Fabio Ciotti relativa a ciò che dovrebbe essere una ricerca in ambito digitale è possibile recuperare termini o coppie di termini che appartengono anche al lessico della LC:

<digital form>, <computational analysis methods, tools and frameworks>, <visualization tools>, <archive of documents>, <platform>, <ontology>, <standards compliance>, <interoperability>, <digital resource the status of research product>

*When I talk about digital research products obviously I'm not referring to the classic article or monograph (although in digital form). Anyone who has really been involved in a research project in Digital Humanities knows that most of the research work consists in the modeling and creation of the data sets, computational analysis methods, tools and frameworks. If the goal of a research project is the creation of an archive of documents, a digital scholarly edition or a visualization tools to display historical data, those are the real products of research, not the articles describing them (often tedious and report style). The assessment of Digital Humanities research must apply to these objects. The creation of a platform of this kind requires: an ontology of the roles and functions performed by the researchers involved in a project;- a set of minimum quality requirements (standards compliance, interoperability, scientific control, coverage of the domain, etc.) to assign a digital resource the status of research product;- a metric based on these requirements to enable* (Ciotti, 2015) [11].

---

<sup>1</sup> Per la prima volta in una pubblicazione ufficiale venne inserita l'espressione Computational Linguistics nel famoso ALPAC Report: *Language and Machines Computers in Translation and Linguistics. National Research Council Washington, D. C. 1966* [Capitolo Automatic Language Processing and Computational Linguistics, pagina 29-31]

## OPERE CITATE

- [1] Domenico Fiormonte, Teresa Numerico, Francesca Tomasi (2015). *The Digital Humanist: A Critical Inquiry*. punctum books, Brooklyn, New York. Pagina 18.
- [2] Anthony Kenny, *Introduction*, in Jean-Philippe Genet, Antonio Zampolli (a cura di) (1992). *Computers and the Humanities* European Science Foundation, Dartmouth. Pagina xii.
  - a. Anthony Kenny, *Introduction*, in Jean-Philippe Genet, Antonio Zampolli (a cura di) (1992). *Computers and the Humanities* European Science Foundation, Dartmouth. Pagina xiii.
- [3] Terras Melissa (2016). A Decade in Digital Humanities. *Journal of Siberian Federal University. Humanities & Social Sciences*. 9. 1637-1650. Pagina 1637
  - a. *Sciences*. 9. 1637-1650. Pagina 1637
- [4] *Actes du Colloque International sur la Méchanisation des Recherches Lexicologiques*. Besançon, Juin (1961). Cahiers de Lexicologie, Vol. 3, 61, Publiés par B. Quemada. Pagina 18.
- [5] Stindlová Jitka, Mater Erich (Rédaction). (1968). *Les machines dans la linguistique: colloque international sur la mécanisation et l'automation des recherches linguistiques*, Prague. Academie Tchecoslovaque des Sciences et 'Academie Allemande des Sciences.
- [6] *Literary Data Processing Conference Proceedings* (1964). IBM Corporation, New York.
- [7] Dom J Froger (1968]. *La critique des textes et son automation*. Dunod, Paris, 1968. Pagina XIII.
  - a. Dom J Froger (1968]. *La critique des textes et son automation*. Dunod, Paris, 1968. Pagina VI.
- [8] Emmanuel G. Mesthene, Technology and humanistic values. *Computers and the Humanities*, 1969, Volume 4, Number 1, Page 1.
- [9] Christine A. Montgomery, Linguistics and Automated Language Processing *International Conference on Computational Linguistics Coling 1969*: Preprint No. 41, ACL. Pagina 2-3.
- [10] Thomas H. Ohlgren, The first international conference on automatic processing of art history data and documents: A report. *Computers and the Humanities*, 1980, Volume 14, Number 2, Pagina 113-114.
- [11] Fabio Ciotti, *A landscape of Digital Humanities in Italy: backgrounds, institutions and infrastructures*, Testo & Senso, n. 16, 2015. [www.testoesenso.com](http://www.testoesenso.com).
- [12] Louis T. Milic, Prospect, *Computers and the Humanities*, 1966. Volume 1, Number 1, Page 1.
- [13] Bernard Wishy , New hardware for the humanities. *Computers and the Humanities*, 1967, Volume 2, Number 1, Page 1.
- [14] Jacques Petit (1961). II. Histoire Littéraire. In *Actes du Colloque International sur la Méchanisation des Recherches Lexicologiques*. Pagina 169.
- [15] Roberto Busa S.J., Antonio Zampolli (1968). Centre pour l'Automation de l'Analyse Linguistique (C:A:A:L.), Gallarate. In Stindlová Jitka, Mater Erich (Rédaction). (1968). *Les machines dans la linguistique: colloque international sur la mécanisation et l'automation des recherches linguistiques*, Prague. Academie Tchecoslovaque des Sciences et 'Academie Allemande des Sciences. Pagina 26.
- [16] Computer Programs Designed to Solve Humanistic Problems (1966). *Computers and the Humanities*, Volume 2 (1966).

# Time machines for Online Services: An Evaluation of a New Interface to Visualize Knowledge Over Time

James Blustein, Nicola Raffaele Di Matteo  
Dalhousie University, Halifax NS

## INTRODUCTION

- Tools to visualize knowledge and facts stored in repositories online use text structures, taxonomies, and image to represent the data organization to human beings.
- Appropriate interfaces could help users to compare data over time, navigate through time easily, find references within the past and make inferences about the future.
- The interface becomes a time machine, a structure that improves the representation of the knowledge.
- New tools to navigate over time have been proposed in the literature.
- In a recent paper, we have introduced a 3-dimensional interface (Di Matteo & Blustein, 2018) that is inspired by technologies used to visualize and study trajectories of atomic particles (Wolter et al., 2009), and NASDAQ data (Kenny & Becker, 2015).
- Users could have the sense of the evolution in time of physical structures, paths, and entities, such as changes in cities and population.

## RESEARCH PROPOSAL

An evaluation of the usability and the efficiency of the proposed interface to navigate over time

A method to compare the interface proposed with those that are the most used in online services using elements such as tables, and slides and buttons

A high fidelity prototype of the interface to navigate the content extracted from the Europeana digital platform (Europeana Pro. History, n.d.) is suggested

## PROPOSED RESEARCH METHODOLOGY

- Informal research questions such as:
  - "Can the user describes facts happened in the past after having only used the interface?"
  - "Is a fact positioned in the past reachable in less time than using traditional solutions?"
  - "Are the functionalities of the interfaces easy to use?"
- A high fidelity prototype
- Usability, efficiency, and satisfaction of the users will be determined using methods of the User-centered design process
- The experiment should point out that the usability of the interface is different from the others considered

## BUILDING THE PROTOTYPE

- It will visualize data from a vast archive online such as Europeana (Europeana Collection, n.d.).
- Europeana gives the access to contents of museums, galleries and audiovisual archives.

- The interface will give a visualization to navigate over time associated elements of interest.
- Starting from 2-dimensional objects such as maps (Europeana Maps, n.d.) available on Europeana, a representation of the places in 3D can be created.
- Europeana offers several APIs to access its objects (Virtuoso SPARQL Editor, n.d.; Europeana SPARQL API, n.d.).
- The high fidelity prototype will present only some fictitious contents and features.

## **METHODS TO GATHER USABILITY, EFFICIENCY AND USER SATISFACTION**

- Structured and contextual interviews in different sessions
- Contextual inquiry to observe how the users interact with the interface
- Questions for an initial exploration of the behaviors of the users, gathering the interesting features, measuring the satisfaction
- Participants should be randomly chosen
- Open questions could be added to gather opinions and reaction in more details

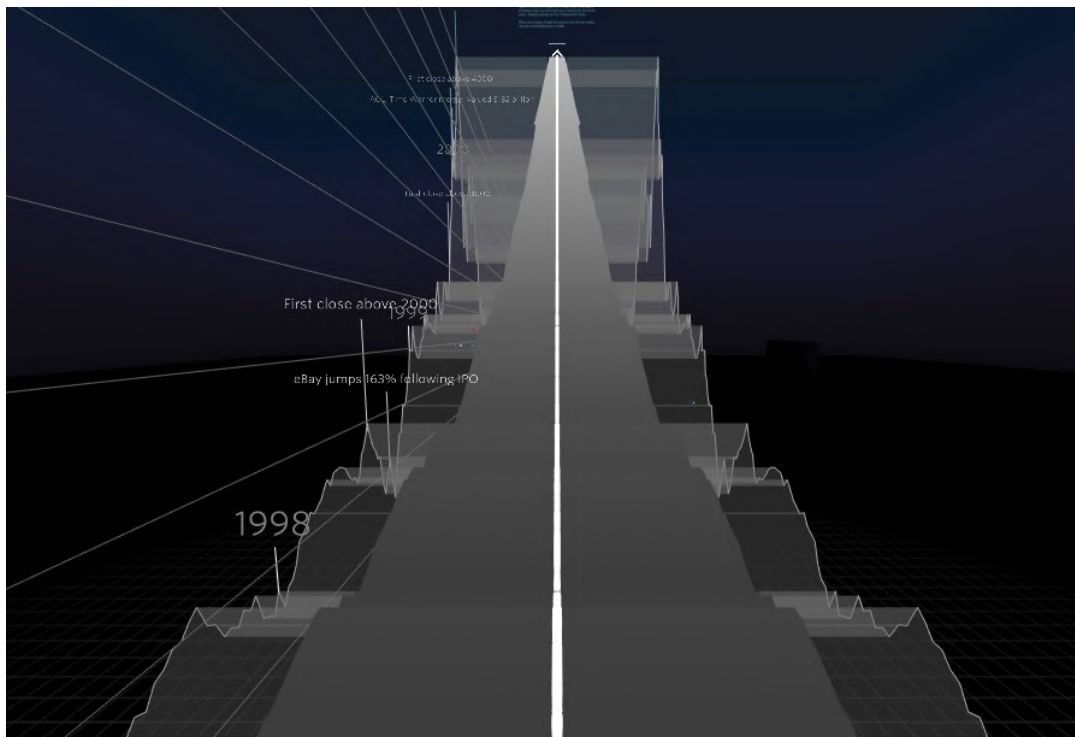


Figure: A virtual reality guided tour of 21 years of the NASDAQ (Kenny & Becker, 2015)

## **ANALYSIS OF THE STUDY DESIGN**

- Research questions are particularly difficult to answer because they involve the use of features that are complex.
- Use cases to test for the most salient parts of the interface
- Unconscious behaviors can be gathered observing how the participants use the prototype with contextual inquiries.
- Results could be compared between groups using statistical models.

## REFERENCES

- [1] Di Matteo, N. R., & Blustein, J. (2018). Navigating data over time: Including the fourth dimension in visualizing knowledge. In *Proceedings of the 1st workshop on human factors in hypertext* (pp. 5–8). doi: 10.1145/3215611.3215614
- [2] *Europeana collection*. (n.d.). Europeana Foundation.
- [3] *Europeana maps*. (n.d.). Europeana Foundation. *Europeana pro. history*. (n.d.). Europeana Foundation. *Europeana sparql api*. (n.d.). Europeana Foundation.
- [4] Kenny, R., & Becker, A. A. (2015). Is the NASDAQ in another bubble. *A virtual reality guided tour of 21 years of the Nasdaq*.
- [5] *Virtuoso sparql editor*. (n.d.). Europeana Foundation.
- [6] Wolter, M., Hentschel, B., Tedjo-Palczynski, I., & Kuhlen, T. (2009). A direct manipulation interface for time navigation in scientific visualizations. In *3dui* (pp. 11–18). doi: 10.1109/3DUI.2009.4811199



# Uno sguardo nei depositi: l'esposizione multimediale dei reperti ceramici di Castelnovo del Friuli

Irene Sarcinelli

Raccolta archeologica di Villa Sulis, "La ceramica di Castelnovo" (Castelnovo del Friuli, PN)  
Univerza na Primorskem/Università del Litorale (Koper/Capodistria)  
irene.sarcinelli@live.com

## ABSTRACT

Il Progetto coinvolge la piccola Raccolta archeologica di Villa Sulis (Castelnovo del Friuli, PN), che espone poco più di un centinaio degli oltre 5000 reperti di ceramica, databili tra metà e fine XVI secolo, rinvenuti in territorio castellano. I materiali non esposti sono attualmente conservati nei magazzini del MAN di Aquileia e non sono pertanto fruibili liberamente da parte del pubblico.

Per valorizzare e promuovere questa piccola realtà museale, si vuole realizzare un'esposizione multimediale di una selezione di reperti tra quelli conservati in magazzino, che amplierà e integrerà l'attuale percorso di visita, restituendo alla comunità parte inedita del suo patrimonio culturale.

Si genereranno 15 modelli tridimensionali di frammenti ceramici non ancora pubblicati utilizzando il software 3DF Zephyr, i quali verranno importati in ambiente Unity 3D per produrre un'esperienza di visita virtuale immersiva e interattiva. L'utente potrà muoversi attraverso la riproduzione di alcune sale museali, interagendo con i reperti esposti guidato da una voce narrante e da una serie di pulsanti, che condurranno a scene di approfondimento dei temi trattati. L'esperienza virtuale è concepita per impegnare l'utente per circa 10-15 minuti e sarà veicolata da una postazione dedicata, in museo.

Grazie alla versatilità di Unity, anche l'esperienza di visita tradizionale sarà aumentata: un'applicazione AR, fruibile da dispositivo mobile fornito in loco, consentirà all'utente di accedere a contenuti aggiuntivi quali didascalie estese con o senza voce narrante, brevi video di storytelling, modelli 3D di oggetti non esposti.

L'intero progetto è stato concepito per essere attuato con poche risorse: utilizzando software a licenza gratuita, hardware proporzionato alle necessità e personale formato nell'umanistica digitale, si punta a ottenere prodotti scientificamente validi, versatili e poco onerosi da gestire a breve e lungo termine. Metodi e strumenti possono essere impiegati per realtà museali simili o in ottica collaborativa, per creare percorsi espositivi inter-museali.

## PAROLE CHIAVE

Realtà virtuale; realtà aumentata; Unity 3D; fotogrammetria; valorizzazione piccoli musei; ceramica post-medievale

## PROPOSTA POSTER

Il Progetto coinvolge la piccola Raccolta archeologica di Villa Sulis (Castelnovo del Friuli, PN), che espone poco più di un centinaio degli oltre 5000 reperti di ceramica, risalenti al periodo compreso tra la metà e la fine del XVI secolo, rinvenuti in territorio castellano. Il corpus ceramico è composto da scarti di fabbricazione, oggetti allo stato finito o semilavorato prodotti dell'attività di botteghe artigianali: si tratta soprattutto di forme aperte

(ciotole/scodelle in particolare) in ceramica dipinta, graffita e rivestita (a ingobbio e/o vetrina), in mono- o policromia, il cui repertorio dei motivi ornamentali è riconducibile a quello della fase tarda della “graffita arcaica”. I reperti costituiscono testimonianza di un contesto produttivo al momento ancora non definibile per scala di produzione e coordinate areali precise, dato che in fase di scavo non sono state individuate tracce di strutture produttive, ma soltanto gli elementi che consentono di riconoscere nel deposito di Castelnovo un probabile scarico di materiale, rappresentativo di una produzione a carattere locale, in serie. Lo studio dei materiali e dei documenti d’archivio suggerisce che a Castelnovo esistessero delle officine ceramiche specializzate nella produzione di stoviglie e recipienti a uso domestico legati, in particolare, alla mensa. I materiali recuperati in fase di scavo ma non confluiti nell’esposizione di Villa Sulis, si trovano attualmente in deposito nei magazzini del MAN di Aquileia e non sono dunque fruibili da parte del pubblico.

La volontà dell’Amministrazione comunale e della direzione della Raccolta archeologica è quella di restituire alla comunità una parte del suo altrimenti non accessibile patrimonio culturale e rafforzarne il potenziale attrattivo, soprattutto verso i giovani e i residenti, per garantire alla Raccolta di essere orgoglio identitario della comunità e di riceverne cura.

Considerata l’esiguità dello spazio disponibile a Villa Sulis per l’installazione di nuove teche e/o pannelli informativi e tenuto conto del rapporto costi/benefici di queste operazioni, si è ritenuto di utilizzare come strumenti privilegiati per il raggiungimento di tali obiettivi le tecnologie digitali.

Tra le premesse metodologiche del progetto vi è una seria riflessione sull’effettivo ruolo che gli strumenti informatici e digitali possono svolgere nella divulgazione scientifica: si è infatti consapevoli del fatto che la computer-based visualisation sia un’approssimazione e un’interpretazione del bene culturale coinvolto. Ci si propone di progettare la relazione tra bene culturale e utente insieme al processo tramite cui questa relazione avviene, qui mediata da strumenti digitali che hanno il compito di avvicinare e servire il visitatore, aiutandolo ad andare oltre ciò che è esposto.

Cuore del progetto è la realizzazione di un’esposizione multimediale di una selezione di pezzi tra quelli conservati in magazzino, che amplierà e integrerà l’attuale percorso di visita.

Si genereranno 15 modelli tridimensionali di frammenti ceramici inediti utilizzando il software 3DF Zephyr (fotogrammetria automatica digitale), i quali verranno importati in ambiente Unity 3D (cross-platform game engine) per produrre un’esperienza di visita virtuale immersiva e interattiva. L’utente potrà muoversi attraverso la riproduzione di alcune sale museali interagendo con i reperti esposti, ingrandendoli, ruotandoli, osservandone dei particolari, guidato da una voce narrante e da pulsanti che condurranno a scene di approfondimento dei temi trattati. L’esperienza virtuale è concepita per impegnare l’utente per circa 10-15 minuti e sarà veicolata da una postazione dedicata, all’interno di Villa Sulis (provvista di tablet e visore per la realtà virtuale), per consentire a chiunque di accedervi sotto la guida dello staff della Raccolta, indipendentemente dal dispositivo personale e dalle competenze digitali posseduti.

Grazie alla versatilità di Unity e alla disponibilità di Plug-in appositi, anche l’esperienza di visita tradizionale sarà aumentata: un’applicazione AR, fruibile da dispositivo mobile fornito in loco, consentirà all’utente di accedere a contenuti aggiuntivi quali didascalie estese con o senza voce narrante per una selezione di reperti (con indicazione di: tipologia, misure, tecnica di realizzazione, principali difetti di fabbricazione riscontrati sul corpo ceramico), mappe, disegni, brevi video di *storytelling* (1-2 minuti), modelli 3D di oggetti simili non esposti.

Vista la natura della Raccolta archeologica, l’esposizione virtuale e l’esperienza AR avranno come filo conduttore l’esplorazione delle tecniche di produzione ceramica e la comparazione di stile e tecnica di realizzazione dei motivi decorativi. Sarà poi possibile procedere al rinnovamento dei contenuti dell’applicazione e dell’esposizione virtuale periodicamente,

mantenendo intatte interfaccia e carattere dell'esperienza. Entrambe le esperienze, inoltre, saranno fruibili in italiano e in inglese.

Grazie al supporto dell'Associazione Scuola per la Ceramica di Villa Sulis, che ha sede in loco, verranno organizzati dei laboratori didattici per bambini e ragazzi, che in tal modo potranno completare l'esperienza di visita museale sperimentando la componente materiale di ciò che hanno imparato a conoscere in ambiente virtuale.

Per la promozione del progetto e dei suoi risultati, considerando il target di riferimento, ci si affiderà alle ormai consolidate tecniche dello *storytelling* via social network, attraverso i canali di comunicazione già attivi per la raccolta archeologica.

Il progetto è pensato per essere completato entro l'inverno 2019: attualmente è pronta la versione dimostrativa dell'applicazione che consentirà la visita in realtà aumentata alla raccolta archeologica. Responsabile del progetto è la direttrice della Raccolta, archeologa diplomata in *Digital humanities* con esperienza di rilievo fotogrammetrico e sviluppo in Unity 3D, che si avvarrà, nella fase intermedia e in quella finale, della consulenza di un programmatore professionista.

L'impiego di *game engine* come Unity 3D per la comunicazione del patrimonio culturale è noto nella letteratura scientifica e non mancano esempi di esposizioni virtuali e applicazioni AR già pubblicati. Attualmente, i riferimenti metodologici internazionali principali nel campo della *computer-based visualisation* del patrimonio culturale sono *The Seville Principles* e *The London Charter*, di cui si è voluto tenere conto per definire metodi coerenti con gli obiettivi del progetto.

Si vuole sottolineare come l'intero progetto sia stato concepito per valorizzare una piccola realtà culturale, che non dispone di molte risorse: utilizzando software a licenza gratuita, hardware proporzionato alle necessità e personale formato nel campo dell'umanistica digitale, si punta a ottenere prodotti scientificamente validi, versatili e poco onerosi da gestire a breve e lungo termine. Tali metodi e strumenti possono essere impiegati anche per altre realtà museali simili o in ottica collaborativa, per creare percorsi espositivi inter-museali.

## RINGRAZIAMENTI

Si ringraziano: il Comune di Castelnovo del Friuli; la Soprintendenza Archeologia, Belle Arti e Paesaggio del Friuli Venezia-Giulia; L'Associazione Scuola per la Ceramica di Villa Sulis; la dott.ssa Elisa Sartori.

## BIBLIOGRAFIA

- [1] Anderson, E. F., Mcloughlin L., Liarokapis F., Peters C., Panagiotis P., De Freitas S. 2010. *Developing serious games for cultural heritage: A state-of-the-art review*. In: *Virtual Reality* 14, 4, pp. 255-75.
- [2] Ch'ng, E. Gaffney, V., Chapman, H. (a cura di). 2013. *Visual Heritage in the Digital Age*. Londra: Springer-Verlag.
- [3] Champion, E. 2011. *Playing with the Past*. Londra: Springer-Verlag.
- [4] Ghouaiel, N., Cieutat, J.-M., Jessel, J.-P., Garbaya, S. 2017. *Mobile Augmented Reality in Museums: Towards Enhancing Visitor's Learning Experience*. In: *The International Journal of Virtual Reality*, 2016, 17(01), pp. 21-31.
- [5] Geronikolakis, E., Tsioumas M., Bertrand S., Loupas A., Zikas P., Papagiannakis, G. 2018. *New Cross/Augmented Reality Experiences for the Virtual Museums of the Future*. In: *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection, Proceedings of the 7th International Conference, EuroMed 2018 (Nicosia, Cyprus, October 29 - November 3, 2018), Part II*. Springer Nature, pp. 518-527.

- [6] Kersten, T., Tschirschwitz, F., Deggim, S., Lindstaedt, M. 2018. *Virtual Reality for Cultural Heritage Monuments - from 3D Data Recording to Immersive Visualisation*. In: Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection, Proceedings of the 7th International Conference, EuroMed 2018 (Nicosia, Cyprus, October 29 - November 3, 2018), Part II. Springer Nature, pp. 74-83.
- [7] *London Charter for the computer-based visualisation of cultural heritage*. 2012. Disponibile: <http://www.londoncharter.org/introduction.html>. Accesso al 5 novembre 2018.
- [8] Oikarinen, T. 2016. *Utilisation of a game engine for archeological visualisation*. In: Proceedings of the 43rd Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Volume 1. Oxford: Archaeopress Publishing Ltd, pp. 27-34. Disponibile: <http://archaeopress.com/ArchaeopressShop/Public/download.asp?id={77DEDD4E-DE8F-43A4-B115-ABE0BB038DA7}>. Accesso al 5 novembre 2018.
- [9] Rizvić, S., Tsiafaki, D. 2016. *Visualisation of Cultural Heritage Objects and its Dissemination*. In: COSCH e-Bulletin, No. 3, pp. 1-21. Disponibile: [http://cosch.info/documents/10179/174435/Bull+3\\_COSCH\\_interim+report\\_WG5.pdf/59270740-c549-4a1f-97f9-30bbb8268c9a](http://cosch.info/documents/10179/174435/Bull+3_COSCH_interim+report_WG5.pdf/59270740-c549-4a1f-97f9-30bbb8268c9a). Accesso al 5 novembre 2018.
- [10] *The Seville Principles. International Principles of Virtual Archeology*. 2011. Disponibile: <http://sevilleprinciples.com>. Accesso al 5 novembre 2018.
- [11] Vitri, S., Casadio, P. 2001. *Magistri Scodelari. Produzioni ceramiche a Castelnovo del Friuli nel Cinquecento*. Castelnovo del Friuli: Comune di Castelnovo del Friuli; Roma: Ministero per i Beni e le Attività culturali

# Vita e i libri di Guarnerio d'Artegna: un percorso didattico nel digital storytelling laboratory

Matteo Venier

Università degli Studi di Udine, [matteo.venier@uniud.it](mailto:matteo.venier@uniud.it)

## ABSTRACT

L'esperienza didattica è stata avviata e compiuta nell'a.a. 2017-2018 nell'ambito del Digital Storytelling Laboratory dell'Università di Udine, e ha avuto come ambito tematico la vita e i libri di Guarnerio d'Artegna.

A partire da competenze base di letteratura italiana e di informatica, si è inteso sviluppare le potenzialità dello storytelling digitale in allievi frequentanti la scuola secondaria superiore. In tale ottica sono stati organizzati due laboratori, che hanno coinvolto studenti dei Licei udinesi in alternanza scuola-lavoro; il fine è stato quello di ricostruire la vicenda umana di Guarnerio d'Artegna, insigne figura di bibliofilo attivo in Friuli nel s. XV, il cui lascito librario alla comunità di San Daniele del Friuli ha dato origine all'attuale Biblioteca Civica Guarneriana.

Organizzato di concerto con la Biblioteca Guarneriana stessa, il lavoro è stato supportato, attraverso diverse attività seminari, da alcuni docenti dell'Università di Udine (che hanno introdotto gli allievi ai caratteri fondanti della cultura del s. XV), e da alcuni esperti della Fondazione Memofonte di Firenze (che hanno descritto i principali processi di digitalizzazione del patrimonio culturale, e hanno fornito una prima generale panoramica sull'uso di Wordpress).

La ricerca è stata avviata sulla base di una importante e cospicua bibliografia relativa – basti ricordare qui i lavori di D'Angelo 1970, il catalogo curato da Scalon-D'Angelo-Casarsa 1991, il saggio di Venier 2014 –, ma per lo specifico lavoro di *storytelling* notevole importanza, quale modello di creazione e rielaborazione narrativa di uno specifico contesto culturale rinascimentale-umanistico, ha assunto anche il saggio di Floramo 2015, in particolare la sezione dedicata a Battista da Cingoli, prediletto copista da Guarnerio.

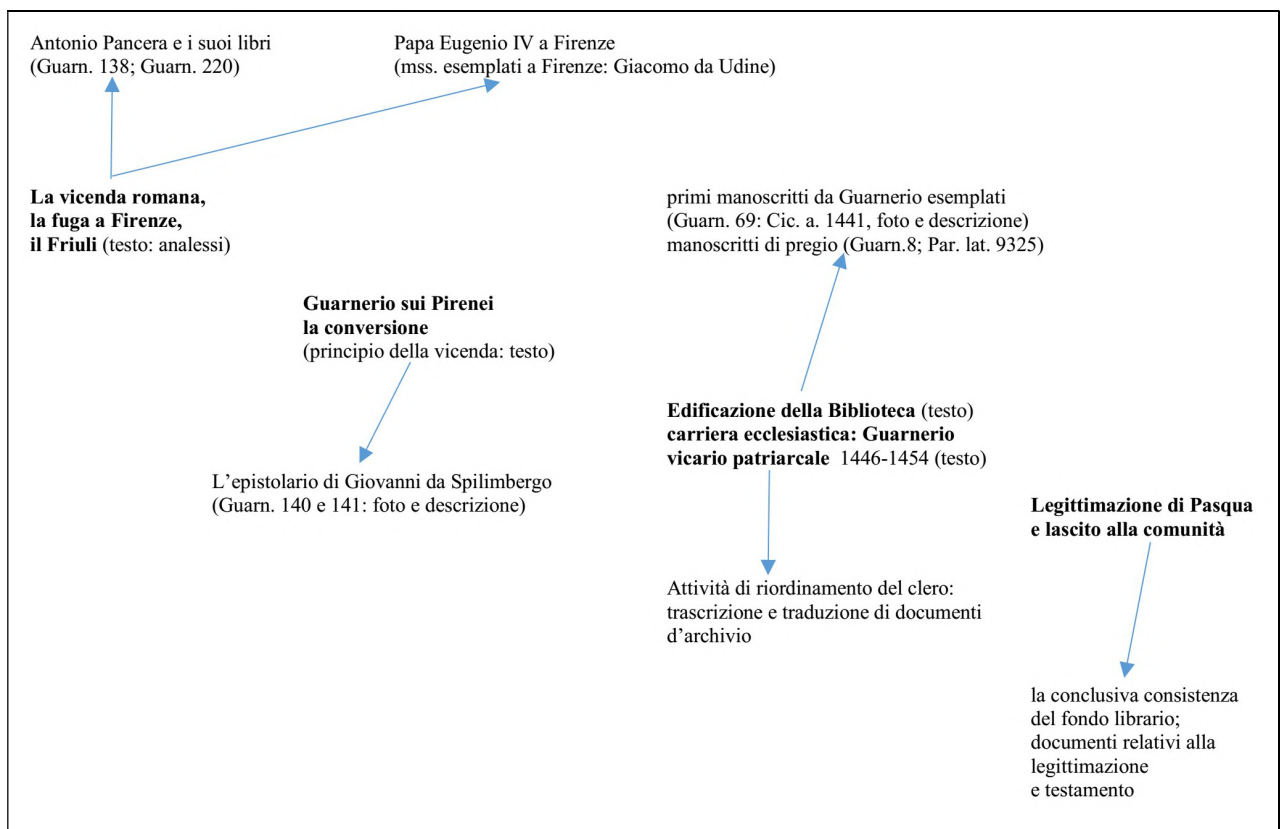
Finalità prima è stata quella di guidare gli studenti, attraverso una didattica hands-on-practice, alla lettura, comprensione, traduzione, commento e quindi digitalizzazione di documenti concernenti la vita di Guarnerio (si tratta di documenti editi in pubblicazioni scientifiche spesso non più accessibili, come il primo inventario della biblioteca); e quindi di condurli, successivamente, al compimento di una narrazione, fondata su quei documenti stessi. Il risultato proviene da un lavoro condiviso: la storia è incardinata su alcuni punti chiave della vicenda dell'umanista: 1. l'apprendistato romano al seguito del cardinal Pancera; 2. il rientro in Friuli; 3. il viaggio sulle Alpi narrato dall'epistolario di Giovanni di Spilimbergo (momento centrale, che segna la 'conversione' del protagonista e la maturazione del suo progetto culturale); 4. la costituzione della biblioteca; 5. il lascito alla comunità.

Coerentemente allo statuto epistemologico dello storytelling, si cercato di adottare una metodologia fondata su collaborazione e inclusività, calando gli apprendimenti in un contesto concreto (la biblioteca di un umanista), funzionale a sviluppare l'autoconsapevolezza, capace, a sua volta, di effettiva «costruzione della conoscenza»: perché sviluppare la competenza narrativa, significa alimentare una risorsa fondamentale per qualsiasi altra disciplina: nella narrazione infatti il sapere assume i connotati di una esperienza che è attiva e che si definisce progressivamente (De Rossi, Petrucco, 2013). L'apprendimento è stato dunque esito di un lavoro condiviso e cooperativo, basato sull'utilizzo di tecnologie multimediali e digitali.

L'ordinamento e la digitalizzazione di testi e documenti (tratti da archivi e biblioteche) ha avuto come primo termine di confronto e di ispirazione, pur se in un'ottica necessariamente meno strutturata, complessa ed estesa, due opere compiute in anni recenti, e cioè *I libri dei Patriarchi* (<http://www.librideipatriarchi.it/>) e il *Dizionario biografico dei Friulani* (<http://www.dizionario biograficodeifriulani.it/>), cioè due strumenti che hanno per la prima volta reso disponibile in rete un patrimonio di conoscenze e di ricerche, il quale era stato fino a pochi anni prima disponibile solo in formato cartaceo (cfr. rispettivamente *I libri dei Patriarchi*, a cura di C. Scalon, Udine, Deputazione di Storia Patria per il Friuli-Istituto Pio Paschini per la Storia della Chiesa in Friuli, 2014; e *Nuovo Liruti. Dizionario biografico dei Friulani*, a cura di C. Scalon et alii, Udine, Forum, 2006-2011). Anche nel caso del sito dedicato a Guarnerio è stata attuata la modalità di una lettura non solo “sequenziale”, ma anche “ipertestuale”, e ciò per mezzo di *link* realizzati nel testo narrativo (o espositivo), ad evidenziare i collegamenti più interessanti, e che sono stati la fonte della narrazione.

La possibilità di creare collegamenti ipertestuali ha reso il lavoro cooperativo fruttuoso, sia da un punto di vista didattico (perché gli allievi sono stati suddivisi in gruppi, che avevano il compito di lavorare a temi diversi, ma fra loro interconnessi, e che sarebbero stati quindi collegati e ricomposti al momento della realizzazione del sito), sia dal punto di vista strettamente scientifico: la sinergia prodotta fra testi studiati e immagini dei documenti oggetto di ricerca (immagini appositamente ricavate con le attrezzature del DSLab presso la Biblioteca Guarneriana), ha permesso di ottenere risultati originali, anche rispetto a testimonianze già in passato pubblicate ma non sempre correttamente interpretate (è il caso della lunga postilla contenuta nel ms. Guarner. 54, firmata da Bartolomeo Baldana, amico di Guarnerio, e in genere ascritta al periodo del suo soggiorno romano, ma in realtà sicuramente da posticipare a un periodo molto più recente, probabilmente agli ultimi anni della vita del Baldana stesso).

Nel progetto, documenti digitalizzati e narrazione avrebbero dovuto confluire in un apposito e autonomo sito, costruito mediante Wordpress, secondo lo schema qui di seguito esposto:



Dove le didascalie in grassetto corrispondono a quelle che avrebbero dovuto essere quattro pagine fondamentali del sito, contenenti i diversi capitoli dello storytelling, cui avrebbero dovuto essere connessi, tramite link, una serie di schede bibliografiche, catalografiche e descrittive, relative ai momenti diversi della vita di Guarnerio.

Senonché nel corso del lavoro è emersa un'urgenza inattesa: prima ancora della realizzazione e pubblicazione del progetto sul web, dato l'esito dell'esperienza e dei materiali raccolti, la Direttrice della Biblioteca Guarneriana ha inteso ospitare l'intero lavoro nel sito già avviato e collaudato della Biblioteca Guarneriana stessa (<http://www.guarneriana.it/>); l'uso effettivo di Wordpress e la mappa sopra descritta sono perciò stati superati, a favore di un'integrazione dei materiali nel sito istituzionale della Guarneriana, il quale adotta un diverso e più complesso sistema (SICAPWeb, cfr. <http://www.cgsi.it/>). In tale prospettiva è stata dunque creata una pagina (base) nel sito del DSLab (<https://storytellinglab.eu/progetto-la-storia-di-guarnerio/>) a supporto del progetto contenente l'introduzione del lavoro svolto e il link al sito della Guarneriana; qui, in modalità integrata, si può visionare il materiale prodotto dal lavoro laboratoriale. I vari materiali (testi e gallerie foto grafiche) sono stati caricati e armonizzati alla struttura del precedente, collaudato sito, dal tecnico del DSLab, Giacomo Vidoni.

Se ciò ha comportato la mancata diretta applicazione delle nozioni concernenti WordPress, che gli allievi avevano appreso grazie ai già citati seminari, peraltro l'integrazione del loro lavoro di ricerca nel sito della Civica Biblioteca di San Daniele ha senz'altro dato all'esperienza il valore aggiuntivo di una virtuosa integrazione fra un'esperienza di didattica e di ricerca in Alternanza scuola-lavoro e una istituzione culturale di riconosciuta rilevanza regionale (e non solo regionale).

Va rilevata finalmente una oggettiva difficoltà nella didattica laboratoriale, relativa alla situazione degli studenti: liceali, come detto, in alternanza scuola-lavoro; tale condizione non è priva di aspetti delicati, sia perché non tutti i partecipanti si sono dimostrati ugualmente motivati, e non tutti hanno frequentato e contribuito con pari serietà; sia soprattutto perché, data la regolamentazione odierna in materia di alternanza, non è previsto un accertamento degli apprendimenti con prove di verifica. Pertanto le effettive ricadute didattiche sono commisurabili in base all'entusiasmo e al coinvolgimento dei partecipanti (che non è mancato), e, soprattutto, in base agli esiti complessivi del prodotto realizzato, che, essendo pubblicato, può essere con oggettività valutato.

## **BIBLIOGRAFIA ESSENZIALE**

- [1] CISOTTO L., *Psicopedagogia e didattica. Processi di insegnamento e di apprendimento*, Roma, Carocci, 2005
- [2] COMAND M., *Digital Storytelling Lab: storie per imparare, per comunicare, per innovare*, in *Disegniamo il futuro. 40 anni dell'Università di Udine*, a cura di A. PINTON e A.F. DE TONI, Udine, Forum, 2018, pp. 213-218
- [3] D'ANGELO M. *Alcune notizie inedite su Guarnerio d'Artegna da un antico regesto*, San Daniele del Friuli, Comune di San Daniele, 1970 («Quaderni guarneriana» 1)
- [4] DE ROSSI M.-PETRUCCO C., *Le narrazioni digitali per l'educazione e la formazione*, Roma, Carocci, 2013
- [5] FLORAMO, *Guarneriana segreta*, Udine, BEE, Bottega errante edizioni, 2015
- [6] SCALON C.-D'Angelo M.-Casarsa L., *La libreria di Guarnerio d'Artegna*, Udine, Casamassima, 1991
- [7] VENIER M., *Guarnerio e i libri degli umanisti*, in *I libri dei patriarchi*, a cura di C. SCALON, Udine, 2014, pp. 307-352

**SITOGRAFIA:**

<http://www.librideipatriarchi.it/>

<http://www.dizionariobiograficodeifriulani.it/>

<http://www.guarneriana.it/>

<https://storytellinglab.eu/progetto-la-storia-di-guarnerio/>



# A Catalog of Web App for Smart Teaching

Marco Corbatto, Antonina Dattolo

SASWEB Research Lab, Department of Mathematics, Computer Science, and Physics, University of Udine, Italy

marco.corbatto@uniud.it - antonina.dattolo@uniud.it

## ABSTRACT

AppInventory is a Web platform, which proposes a holistic view on a digital catalog of 271 applications, an innovative visual layout, and semantic browsing structures; the applications have been organized in an original taxonomy and can support users in creating and sharing digital artifacts, aggregating, remixing and collecting heterogeneous materials and communicating within working groups.

## KEYWORDS

Web 2.0 applications repository, App 2.0 taxonomy, Multimedia design and development for smart e-learning, Innovative smart teaching and learning technologies, Multimedia for user engagement and motivation in education.

## INTRODUCTION

The Web offers hundreds of applications, which a challenging opportunity for teachers who would like to experiment and adopt student-centred methodologies and use them into daily teaching and learning activities; opportunely used, these applications can improve the collaborative, cognitive and creative work of the students, and enhance and redefine traditional educational practices. Nevertheless, although these applications are generally easy to find and use, it is often difficult, for a teacher, to identify the right one for a specific task, and to have a general awareness on their availability their potential in an educational setting.

AppInventory [5, 6] proposes a new platform and an innovative browsing tool for overcoming these limitations. The Web platform contains a multimedia catalog of 271 applications, organized by applying an original taxonomy, offering users the opportunity to explore according to personal paths, assigning ratings and leaving comments to the applications. The AppInventory platform is consistent with the objectives of the European Digital Competences Framework for Citizens 2.0 (DigiComp) [13]: in particular, it can contribute to the development of ten of the twenty-one competence dimensions stated in the DigiComps conceptual reference model.

## RELATED WORK

Several repositories index applications, proposing classification and evaluation [2, 3, 11] schemes.

A rich discovery engine of websites, mobile apps, desktop programs, and electronic products for teaching and learning is EdShelf [9]: a user can filter the tools by price, platform, subject, age, category and keywords. Unfortunately, subject and category are two long flat lists of keywords. Essediquadro [10] is a service of documentation and orientation on the teaching software and on other resources for the learning process. The tools can be searched by subject of study (Mathematics, Italian, etc.) and by specific subject matter, but the category of the tools is not considered. Similar search fields are proposed by Apps4edu [1]. It is possible to list all the apps in it, but the result is a flat, unusable, paged-list of tools. CSE (Common sense

education) [4] invites one to find the 'perfect tool'; it introduces the interesting, abstract concept of purpose, but it is used more as teaching context of use more than real purpose.

## OPEN ISSUES AND OUR AIMS

Related work highlighted some open challenges and weaknesses:

- the navigation and searching of tools do not offer a general overview, but long lists of applications, often difficult to read;
- there is a complete lack of graphic views, which could offer users a visual, holistic idea of the existing tools;
- the concept of category as a tool is often thought of as a subject of study, or context of use and not as purpose for teachers. The existing taxonomies are not purpose-based; the semantic relations among the tools are not highlighted, and the degree of belonging of a tool to a cluster in the taxonomies is not clear;
- The user interaction is limited.

Our contribution focuses on overcoming these limitations.

## THE WEB PLATFORM

AppInventory is a free Web application (<http://appinventory.uniud.it>), which uses SVG/HTML5/CSS3 Web standards and powered by D3 data driven visualization library, PHP, MySQL server side technologies and AJAX techniques.

Data model is based on zz-structures [7, 8, 12] enriched with specialized zz-views in order to offer original exploration solutions in addition to traditional searches and data filtering. It contains 271 cards of Web 2.0 applications organized into 3 main macro-categories, as shown in Figure 1.



Figure 1. An holistic view of AppInventory.

*Collaboration & Communication tools* contains applications, organized in 7 sub-categories, to manage groups, to collaborate on the same documents online, to support users in planning projects and activities, to interact in real-time on a virtual board or to collect data by surveys and quizzes; *Authoring tools* contains applications (in 13 sub-categories) to support users in

building up digital artifacts of various typologies; *Aggregators* contains applications (in 4 sub-categories), which support users in collecting homogeneous or heterogeneous materials. The size of each circle is proportional to its populosity. Due to the apps' hybrid characterization, the attribution to the categories is multiple and weighted.

**Semantic zoom.** Zooming in additional items (see Figures 2 and 3) become visible at an appropriate zoom level in order to enhance the understanding and minimize the cognitive load. Continue zoom mechanism minimizes users disorientation.

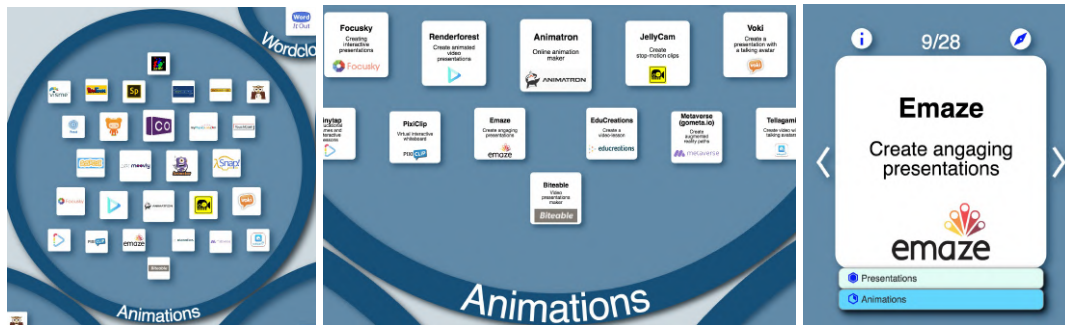


Figure 2. Views for different zooming levels.

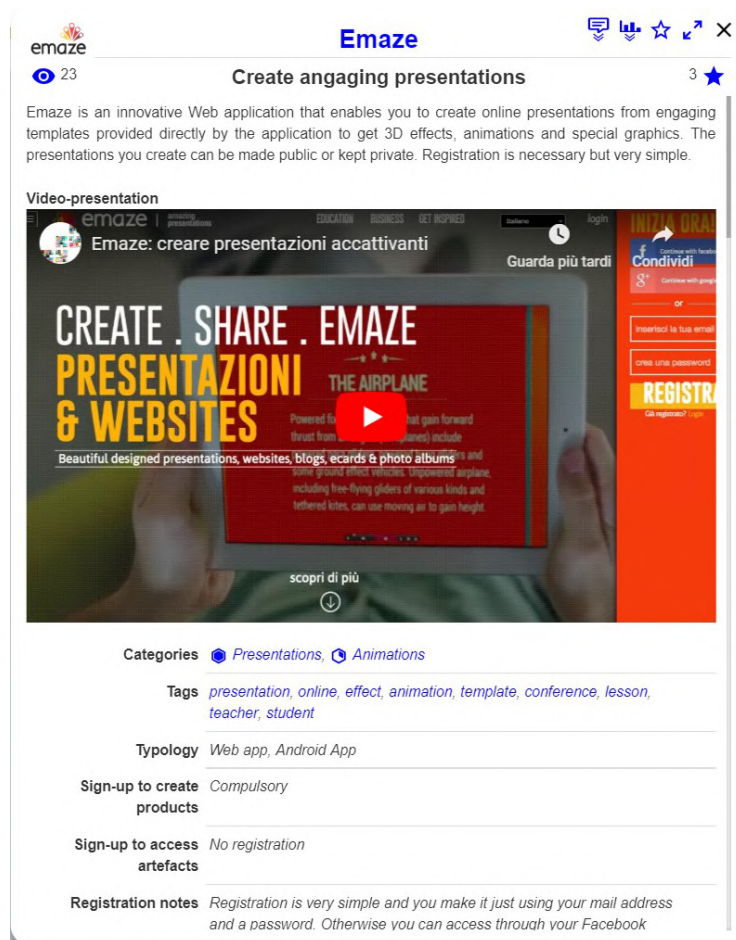


Figure 3. The card of the Emaze app.

Navigation mechanisms. Based on zz-structure's ranks, following Figure 4:

- Left/right arrows move to previous/next app in the current navigation set of applications (default: apps in the same category). Navigation window shows the list of the apps in the current navigation set.
- Select a new navigation set by clicking on the compass on top-right.
- Each criterion is marked with a color.
- Criteria can be combined with AND/OR operators; compound criteria are marked in yellow.

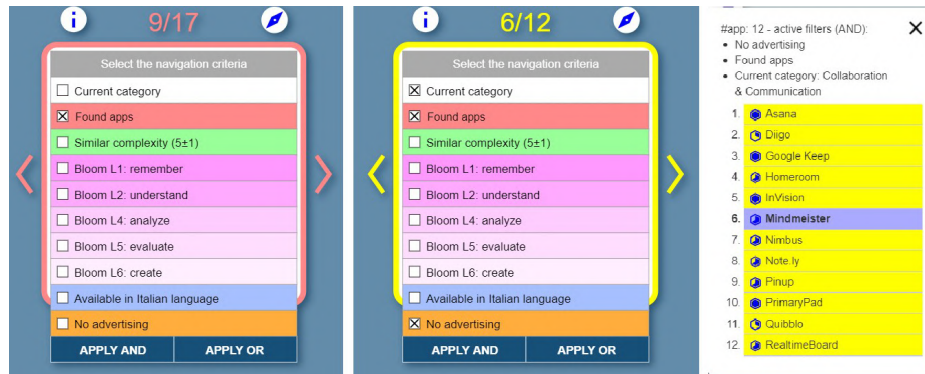


Figure 4. Navigation mechanisms.

**App ratings.** AppInventory collects user ratings about the apps on four aspects plus a general one, as shown in Figure 5.



Figure 5. Ratings.

**User contribution.** AppInventory collects users' contributions for the single app: comments; reports on description card; descriptions of original use cases and for the overall catalog: suggestions of new apps.

## CONCLUSIONS AND FUTURE WORK

More details on AppInventory is available on the dedicated Website. Future work will involve the constant updating of the information published in the catalog, the development of new views and new features, the optimization of the Web interface for mobile devices, the experimentation of a recommender system based on zz-structures and a more complete user evaluation which allows to consider different aspects of the platform and the effectiveness of the adopted solutions

## ACKNOWLEDGMENTS

We would like to thank all students that participated to this project and contributed in different ways to the birth of the catalog. We mentioned all their names on a dedicated page <http://appinventory.uniud.it/en/about-us/> of the AppInventory platform.

## BIBLIOGRAFIA

- [1] Apps4edu. <http://www.uen.org/apps4edu>. Last visited 28 December 2018.
- [2] T. Cherner, J. Dix, and C. Lee, “Cleaning up that mess: A framework for classifying educational apps,” *Contemporary Issues in Technology and Teacher Education*, vol. 14, no. 2, pp. 158–193, 2014.
- [3] T. Cherner, C.-Y. Lee, A. Fegely, and L. Santaniello, A detailed rubric for assessing the quality of teacher resource apps. *Journal of Information Technology Education: Innovations in Practice*, vol. 15, pp. 117-143, 2016.
- [4] Common sense education. <http://www.common sense.org>. Last visited 28 December 2018.
- [5] M. Corbatta, Modeling and developing a learning design system based on graphic organizers, in *Adjunct Publication of the 25th Conf. on User Modeling, Adaptation and Personalization*. ACM, 2017, pp. 117–118.
- [6] M. Corbatta and A. Dattolo, Appinventory: a visual catalogue of web 2.0 and mobile applications for supporting teaching and learning activities. In *Proceedings of the 22nd International Conference Information Visualisation - IV 2018*, Salerno, Italy, July 10-13 2018, pp. 530–535, IEEE.
- [7] A. Dattolo and F. L. Luccio, “A state of art survey on zz-structures,” in *Proceedings of the 1st Workshop on New Forms of Xanalogical Storage and Function*, CEUR, no. 508, Turin, Italy, June 29 2009, pp. 1–6.
- [8] A. Dattolo and F. L. Luccio, “A formal description of zz-structures,” in *Proceedings of the 1st Workshop on New Forms of Xanalogical Storage and Function*, CEUR, no. 508, Turin, Italy, June 29 2009, pp. 7–11.
- [9] Edshelf. <http://edshelf.com>. Last visited 28 December 2018.
- [10] Essediquadro. <http://sd2.itd.cnr.it>. Last visited 28 December 2018.
- [11] A. Jareno, E. M. Morales-Morgado, and F. Martinez, “Design and validation of an instrument to evaluate educational apps and creation of a digital repository,” in *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality*, ser. TEEM '16. New York, NY, USA: ACM, 2016, pp. 611–618.
- [12] T.H.Nelson, “A cosmology for a different computer universe: Data model, mechanisms, virtual machine and visualization infrastructure”, *Journal of Digital Information*, vol. 5, no. 1, July 2004.
- [13] The European Digital Competence Framework for Citizens. <http://ec.europa.eu/jrc/en/digcomp/digital-competence-framework>. Last visited 28 December 2018.