

# GESTIRE I DATI DELLA RICERCA: tutto ciò che c'è da sapere



## LEZIONE 1

Mario Marino

[mario.marino6@unibo.it](mailto:mario.marino6@unibo.it)

Data Steward di Area Sociale, Research Division (ARIC), Alma Mater Studiorum - Università di Bologna



03 Ottobre 2023



## DURANTE L'INCONTRO DI OGGI...

---

- Per favore tenete i microfoni spenti durante la presentazione.
- Sentitevi liberi di accendere microfono e videocamera nei momenti di Q&A.
- Sentitevi liberi di porre le vostre domande in chat in qualsiasi momento, risponderemo durante il Q&A!
- Abbiamo previsto all'interno di questa lezione due pause, che seguiranno i momenti Q&A.

# GESTIRE I DATI DELLA RICERCA: TUTTO CIÒ CHE C'È DA SAPERE

Il dato e il suo  
valore

Open Science

I principi FAIR

Research Data  
Management

Data  
Management  
Plan

---

Modulo formativo diviso in tre incontri.

Unico pre-requisito è avere una propria esperienza di ricerca.



## LE TEMATICHE CHE AFFRONTEREMO OGGI

1

### I DATI DELLA RICERCA

Definizioni e esempi di dominio.

2

### IL VALORE DEI DATI E DELLA LORO CORRETTA GESTIONE

Horror stories e vantaggi associati alla gestione del dato.

3

### OPEN SCIENCE

Il contesto culturale e politico in cui si muove la ricerca.

4

### PRINCIPI FAIR NELLA GESTIONE DEI DATI DELLA RICERCA

Cosa sono e come applicarli nella pratica di ricerca.





# CHI SIAMO? DATA STEWARDS @UNIBO



Il progetto Data Steward @Unibo mira al rafforzamento del supporto alla gestione dei dati della ricerca:

- Chi sono i Data Stewards?

Figure di supporto per le tematiche di gestione FAIR dei dati della ricerca e stesura del Data Management Plan ai team ricerca di UNIBO

- Cosa fanno?

Supportano i ricercatori nella gestione FAIR dei dati (research data management) e nella stesura del Data Management Plan, principalmente nel contesto dei progetti Horizon Europe

Supportano la Governance di Ateneo nella promozione di Open Science

- Qual è il loro profilo?

Background scientifico (esperti di dominio)

Esperienza in data management

FAIR principles e conoscenza di pratiche di OS





MARIO MARINO

Area Sociale



BIANCA GUALANDI

Area Umanistica



GIULIA CALDONI

Area Biomedica

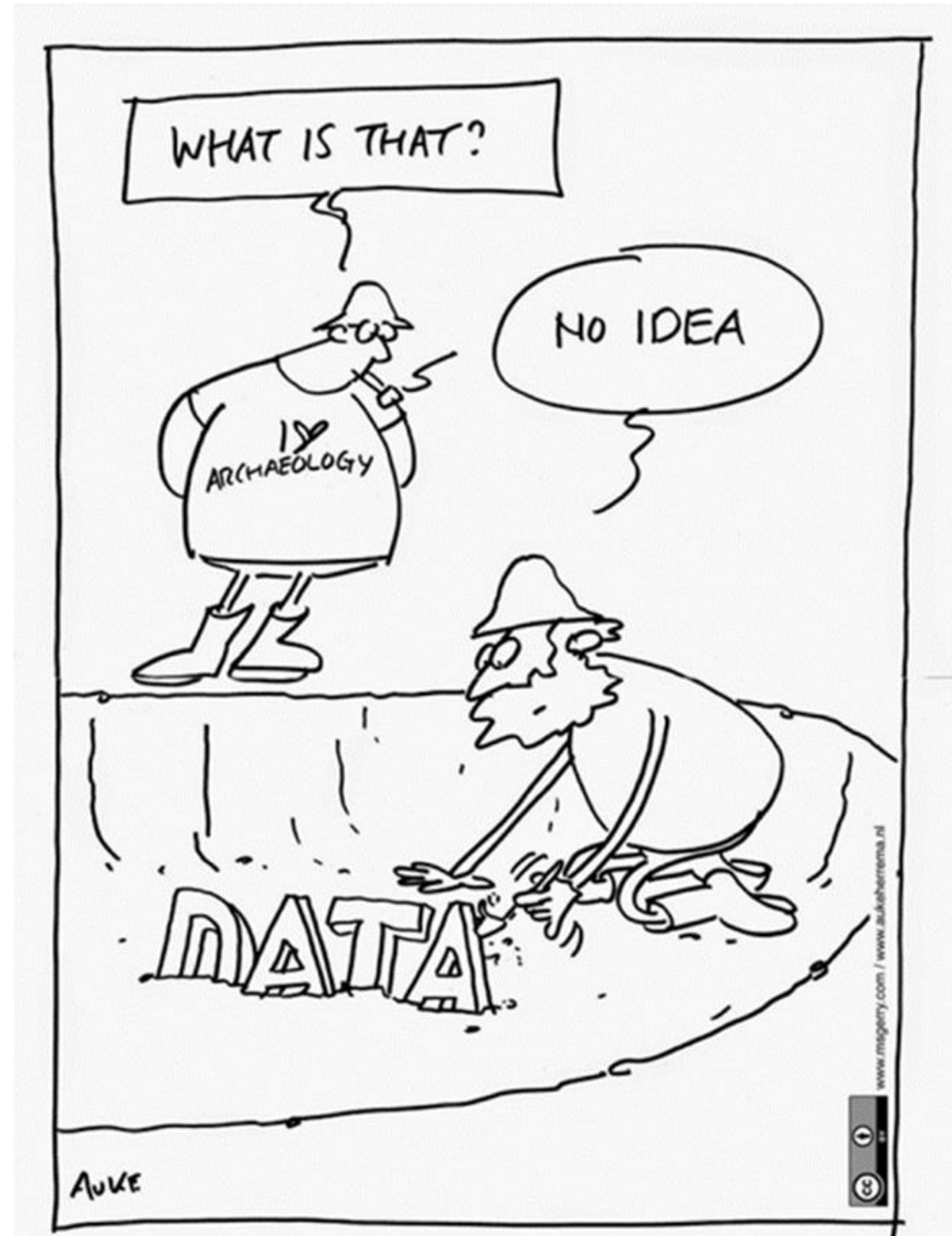


SARA COPPINI

Area Tecnologica

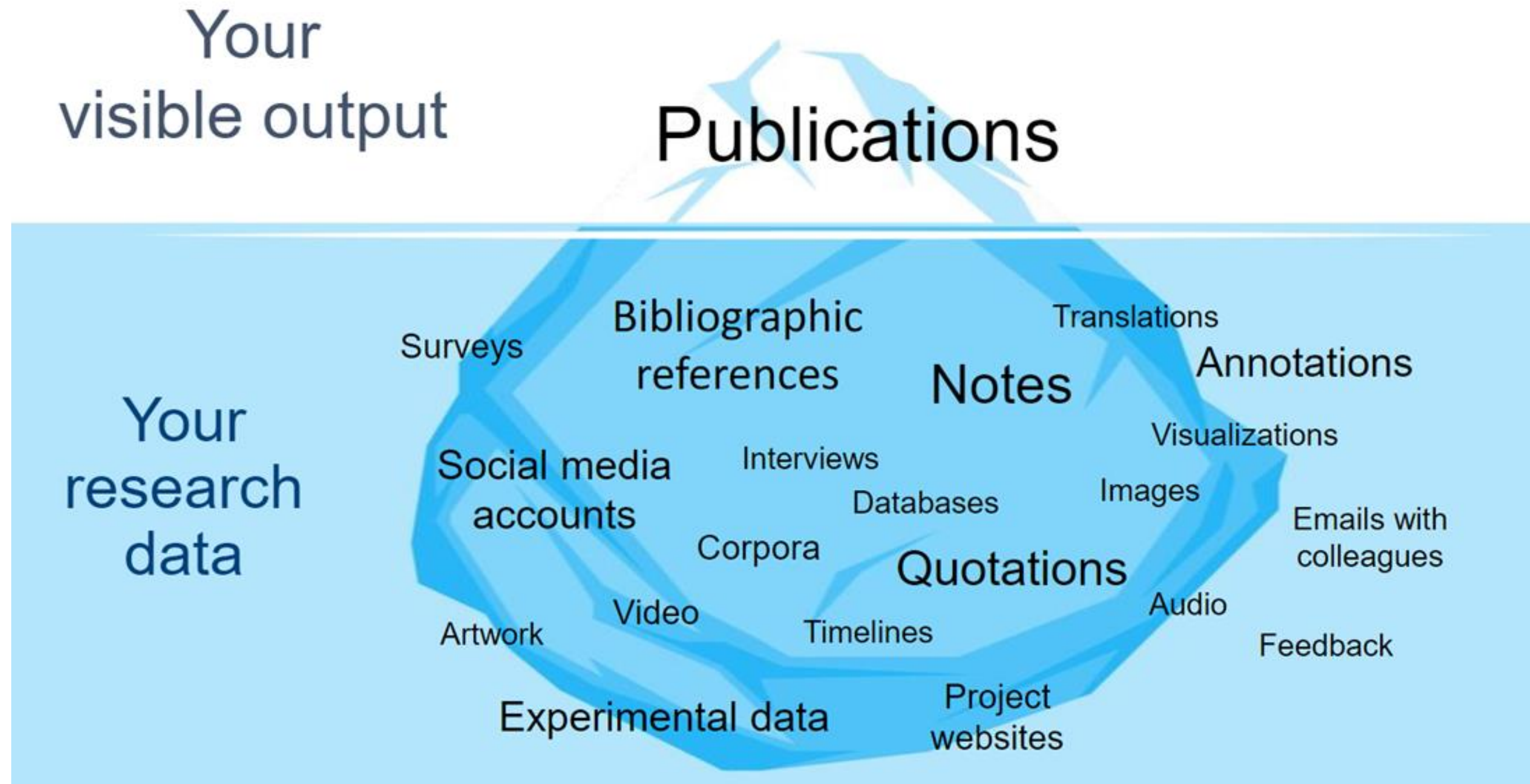


INIZIAMO A PARLARE DI DATI...





## INIZIAMO A PARLARE DI DATI...





# QUALCHE DEFINIZIONE UFFICIALE



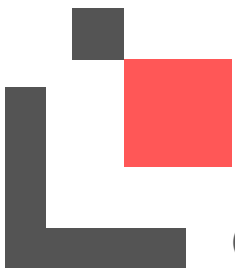


Informazioni, in particolare fatti o numeri, raccolte per essere esaminate e considerate come base per ragionamenti, discussioni o calcoli.





I record fattuali (numeri, testi, immagini e suoni) utilizzati come fonti primarie per la ricerca scientifica e che sono comunemente accettati dalla comunità scientifica come necessari per convalidare i risultati delle ricerche.





L'intersezione tra scienza dell'informazione e tecnologia ha cambiato la natura e il significato del dato: tutto ciò che si può rappresentare sotto forma di bits è dato.







## DIVERSI TIPI DI DATI DELLA RICERCA

Esiste un'enorme varietà di tipi di dati ed è possibile classificarli in modi diversi

- Contenuto: numerico, testuale, audiovisivo, multimediale...
- Formato: fogli di calcolo, database, immagini, mappe, file audio, testo...
- Modalità di raccolta: sperimentale, osservazionale, simulazione, dati derivati/compilati da altre fonti
- Natura digitale (nati digitali o digitalizzati) o non digitale (es. indagini cartacee, appunti...)
- Natura primaria (generati dal ricercatore per un particolare scopo di ricerca o progetto) o secondaria (originariamente creati da qualcun altro per un altro scopo)
- Natura grezza o elaborata



## QUALI TIPI DI DATI POTRESTE INCONTRARE DURANTE LA VOSTRA RICERCA?

Dipende dal vostro ambito di  
ricerca!

Oggi vediamo degli esempi collegati a...

- Database
- Dati tabulari
- Survey
- Mappe
- Algoritmi

IL DATABASE

Un database è una raccolta strutturata e organizzata di dati che sono memorizzati in modo permanente su un sistema di archiviazione, in genere su un computer o un server.

Questi dati sono progettati per essere facilmente accessibili, gestibili e aggiornati.

I database vengono comunemente utilizzati per archiviare informazioni in modo efficiente, consentendo agli utenti di recuperare, manipolare e analizzare i dati in vari modi.

Reddito netto : Regioni e tipo di comune

PersonalizzaEsporta

| Tipo dato                             | reddito medio annuale delle famiglie (in euro) |                 |                                   |        |        |
|---------------------------------------|--|-----------------|-----------------------------------|--------|--------|
|                                       | esclusi fitti imputati                         |                 |                                   |        |        |
|                                       | 2021   |                 |                                   |        |        |
| Presenza affitti imputati             |  |                 |                                   |        |        |
| Selezione periodo                     |  |                 |                                   |        |        |
| Fonte principale di reddito familiare | lavoro dipendente                              | lavoro autonomo | pensioni e trasferimenti pubblici | altro  | totale |
| Territorio                            |  |                 |                                   |        |        |
| Italia                                | 37 414   | 45 392          | 28 134                            | 19 528 | 33 798 |
| Nord-ovest                            | 41 438   | 52 988          | 29 986                            | 26 322 | 37 647 |
| Piemonte                              | 39 258   | 44 609          | 27 568                            | 22 702 | 34 125 |
| Valle d'Aosta / Vallée d'Aoste        | 36 111 (n)                                     | --              | 30 607 (n)                        | --     | 35 237 |
| Liguria                               | 36 288   | 41 462          | 26 107 (n)                        | 26 564 | 31 586 |
| Lombardia                             | 43 017   | 58 864          | 31 993 (n)                        | 28 431 | 40 277 |
| Nord-est                              | 41 297   | 53 459          | 30 915                            | 23 316 | 38 340 |
| Provincia Autonoma Bolzano / Bozen    | 46 368   | 70 485          | 32 803 (n)                        | --     | 44 370 |
| Provincia Autonoma Trento             | 41 750   | 54 715          | 28 044 (n)                        | --     | 38 526 |
| Veneto                                | 39 476   | 51 723          | 30 472 (n)                        | 32 099 | 37 113 |
| Friuli-Venezia Giulia                 | 39 967   | 53 363          | 27 302 (n)                        | --     | 34 637 |
| Emilia-Romagna                        | 42 868   | 53 197          | 32 692 (n)                        | 23 480 | 39 941 |
| Centro                                | 36 850   | 42 190          | 31 156                            | 19 401 | 34 555 |
| Toscana                               | 36 849   | 45 826          | 31 985 (n)                        | 24 009 | 35 815 |
| Umbria                                | 35 998   | 53 150          | 35 092 (n)                        | --     | 38 011 |
| Marche                                | 38 399   | 45 720          | 29 144 (n)                        | --     | 34 730 |
| Lazio                                 | 36 598   | 36 700          | 30 597                            | 16 581 | 33 205 |
| Sud                                   | 31 834   | 33 250          | 23 754                            | 12 761 | 27 467 |
| Abruzzo                               | 27 621   | 27 375          | 23 557 (n)                        | --     | 29 021 |
| Molise                                | 32 827 (n)                                     | 33 576          | 22 620 (n)                        | --     | 27 247 |
| Campania                              | 30 404   | 32 780          | 24 026                            | 12 724 | 26 676 |
| Puglia                                | 32 201   | 34 577          | 22 885 (n)                        | 11 432 | 28 097 |
| Basilicata                            | 31 565 (n)                                     | 49 709          | 22 694 (n)                        | --     | 29 043 |

Legend

Il dato non raggiunge la metà della cifra minima considerata

dato statisticamente non significativo (vedasi metadati di dataset)

Fonte: Istat  
[http://dati.istat.it/viewhtml.aspx?il=blank&vh=0000&vf=0&vcq=1100&graph=0&view-metadata=1&lang=it&QueryId=22919&metadata=DCCV\\_REDNETFAMFONTERED](http://dati.istat.it/viewhtml.aspx?il=blank&vh=0000&vf=0&vcq=1100&graph=0&view-metadata=1&lang=it&QueryId=22919&metadata=DCCV_REDNETFAMFONTERED)

Show Data

| Domain Code | Domain                | Area Code (M49) | Area  | Element Code | Element         | Item Code (CPC) | Item       | Year Code | Year | Unit   | Value | Flag | Flag Desc  |
|-------------|-----------------------|-----------------|-------|--------------|-----------------|-----------------|------------|-----------|------|--------|-------|------|------------|
| FBS         | Food Balances (2010~) | 380             | Italy | 5611         | Import Quantity | F2537           | Sugar beet | 2010      | 2010 | 1000 t | 1     | I    | Imputed va |
| FBS         | Food Balances (2010~) | 380             | Italy | 5611         | Import Quantity | F2537           | Sugar beet | 2011      | 2011 | 1000 t | 1     | I    | Imputed va |
| FBS         | Food Balances (2010~) | 380             | Italy | 5611         | Import Quantity | F2537           | Sugar beet | 2012      | 2012 | 1000 t | 1     | I    | Imputed va |
| FBS         | Food Balances (2010~) | 380             | Italy | 5611         | Import Quantity | F2537           | Sugar beet | 2013      | 2013 | 1000 t | 2     | I    | Imputed va |
| FBS         | Food Balances (2010~) | 380             | Italy | 5611         | Import Quantity | F2537           | Sugar beet | 2014      | 2014 | 1000 t | 2     | I    | Imputed va |
| FBS         | Food Balances (2010~) | 380             | Italy | 5611         | Import Quantity | F2537           | Sugar beet | 2015      | 2015 | 1000 t | 3     | I    | Imputed va |
| FBS         | Food Balances (2010~) | 380             | Italy | 5611         | Import Quantity | F2537           | Sugar beet | 2016      | 2016 | 1000 t | 1     | I    | Imputed va |
| FBS         | Food Balances (2010~) | 380             | Italy | 5611         | Import Quantity | F2537           | Sugar beet | 2017      | 2017 | 1000 t | 0     | I    | Imputed va |

Showing 1 to 22 of 22 rows

Fonte: Faostat  
<https://www.fao.org/faostat/en/#data/FBS>



## I DATI TABULARI

I dati tabulari sono un tipo comune di dati strutturati organizzati in una tabella o griglia, in cui le informazioni sono disposte in righe e colonne. Ogni riga rappresenta una singola osservazione o record, mentre le colonne rappresentano i diversi attributi o variabili associati a ciascuna osservazione. Questa tipologia è ampiamente utilizzata per rappresentare dati in diversi contesti, come l'analisi dei dati, le basi di dati relazionali, i fogli di calcolo e molto altro, dato che presenta numerosi vantaggi:

1. organizzazione in una struttura rigida e ben definita. Ogni colonna ha un nome che rappresenta l'attributo associato, e ogni riga contiene i valori corrispondenti per ciascuna variabile.
2. la struttura tabulare è intuitiva e di facile comprensione, il che facilita la lettura e l'interpretazione dei dati, sia per le persone che per i computer.
3. adatta per l'analisi dei dati, permette di eseguire operazioni come l'ordinamento, il filtraggio, l'aggregazione e il calcolo di statistiche
4. spesso immagazzinati in formati che consentono interrogazioni complesse utilizzando linguaggi come SQL o strumenti di analisi dei dati.

# I DATI TABULARI - FORMATI

## Formati più utilizzati:

- 1. CSV (Comma-Separated Values):** Questo è un formato testuale in cui i dati sono separati da virgole o altri separatori, come il punto e virgola. I file CSV sono ampiamente utilizzati per l'importazione ed esportazione di dati in applicazioni come fogli di calcolo e database.
- 2. Excel:** I fogli di calcolo in formato Excel sono un formato di dati tabulari molto comune. Excel permette di creare, manipolare e analizzare dati tabulari in modo interattivo.
- 3. Database relazionali:** I database relazionali, come MySQL, PostgreSQL e Microsoft SQL Server, utilizzano tabelle per archiviare dati tabulari. Questi database consentono di gestire grandi quantità di dati e supportano query complesse.
- 4. DataFrames:** I DataFrames sono strutture dati tabulari utilizzate in molte librerie di programmazione come pandas in Python o data.frame in R. Sono particolarmente utili per l'analisi dei dati.
- 5. JSON e XML:** Anche se non sono formati tabulari "puri", JSON (JavaScript Object Notation) e XML (eXtensible Markup Language) possono essere utilizzati per rappresentare dati tabulari in modo gerarchico, utilizzando una struttura a oggetti nidificati.

|    | hb030   | DB095     | HX040 | HX050 | PROV | PROV2 | DB030 | DB100 | REGIONE | RT |
|----|---------|-----------|-------|-------|------|-------|-------|-------|---------|----|
| 1  | 8072900 | 846.6803  | 2     | 1.5   | 1    | 1     | 1321  | 1     | 10      |    |
| 2  | 8073200 | 1335.2270 | 3     | 1.5   | 1    | 1     | 16379 | 1     | 10      |    |
| 3  | 8073201 | 2769.2500 | 1     | 1.0   | 1    | 1     | 11846 | 1     | 10      |    |
| 4  | 8073400 | 2797.6740 | 1     | 1.0   | 1    | 1     | 10217 | 1     | 10      |    |
| 5  | 8073500 | 993.2654  | 2     | 1.5   | 1    | 1     | 15216 | 1     | 10      |    |
| 6  | 8073700 | 1080.4490 | 3     | 2.0   | 1    | 1     | 26439 | 1     | 10      |    |
| 7  | 8073800 | 2277.7000 | 1     | 1.0   | 1    | 1     | 1347  | 1     | 10      |    |
| 8  | 8074700 | 1899.7530 | 5     | 2.6   | 1    | 1     | 22466 | 1     | 10      |    |
| 9  | 8075500 | 2245.3150 | 1     | 1.0   | 1    | 1     | 20401 | 1     | 10      |    |
| 10 | 8075700 | 1887.0910 | 1     | 1.0   | 1    | 1     | 7355  | 2     | 10      |    |
| 11 | 8075900 | 1574.7730 | 1     | 1.0   | 1    | 1     | 12527 | 2     | 10      |    |
| 12 | 8076000 | 1199.8730 | 2     | 1.5   | 1    | 1     | 11255 | 2     | 10      |    |
| 13 | 8076200 | 959.3061  | 4     | 2.5   | 1    | 1     | 17818 | 2     | 10      |    |
| 14 | 8076400 | 1102.5950 | 2     | 1.5   | 1    | 1     | 6843  | 2     | 10      |    |
| 15 | 8076700 | 873.7050  | 3     | 1.5   | 1    | 1     | 31304 | 3     | 10      |    |

|    | A  | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|----|--|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1  | PROV2,Anno,name,Quota contribuenti su Popolazione,Reddito medio da Fabbricati,Reddito medio da Lavoro Dipendente,Reddito medio da Pensione,Reddito medio da Lavoro Autonomo, |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 2  | 1,2014,1_2014,0.784749257,1707.848897,22857.10108,18569.7011,39053.41408,21965.58542,-9725.20153,4919.878822,12530.09501,20305.43265,34569.80619,63532.87517,91681.90014,;   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 3  | 7,2014,7_2014,0.824633549,2023.627786,20506.79358,16919.99492,35040.62599,20721.47377,-11221.30337,4709.48558,12579.09707,20368.37084,34417.62422,63619.80873,91079.95328    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 4  | 15,2014,15_2014,0.836213721,1760.117749,27656.21727,20434.94843,49180.8786,26709.59767,-9605.099476,4853.907844,12554.17107,20402.67273,35354.85855,63618.25797,91948.567    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 5  | 25,2014,25_2014,0.828506627,1333.204881,20418.60278,15328.82487,36893.55287,19416.1617,-11800.64407,4778.089189,12544.37684,20205.9067,33621.51408,63549.60016,91500.7250    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 6  | 27,2014,27_2014,0.798524081,1600.112752,20688.74356,16976.30776,39431.93561,20386.71928,-9471.673402,5002.563946,12504.14503,20306.76718,34211.57852,63353.20099,91645.55    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 7  | 29,2014,29_2014,0.815359842,927.0261015,18755.40188,15052.48422,37575.18102,17831.39168,-9905.955186,4982.538724,12523.16452,20033.65505,33640.66439,63335.5557,91899.022    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 8  | 37,2014,37_2014,0.854781308,1914.191566,23711.17684,19047.37152,43054.2342,23412.40586,-9363.734064,5008.238843,12587.5512,20345.02885,34788.56532,63606.26695,91786.9083    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 9  | 41,2014,41_2014,0.794074168,1231.455358,18915.7748,14999.92144,36426.32004,18192.9332,-8882.665076,5076.429048,12442.71078,20037.96938,34077.1196,63484.76586,91403.59129    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 10 | 42,2014,42_2014,0.79987986,1232.979696,19940.21838,16844.62468,34975.33399,19668.6918,-8949.894009,5037.870166,12523.71725,20149.40132,34053.7235,63595.04695,91665.19931    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 11 | 48,2014,48_2014,0.831948896,2227.762166,21601.26149,18049.91143,35404.24706,21760.79132,-10232.67451,4987.427328,12570.91672,20206.59207,34652.63523,63740.38559,91745.27    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 12 | 51,2014,51_2014,0.818065163,1430.84599,19134.35107,16092.84438,32255.94953,18926.04695,-9692.164179,5036.536024,12608.10175,19957.1096,33948.30933,63477.52025,91368.7998    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 13 | 58,2014,58_2014,0.758225207,2495.599105,24057.49173,20672.14447,37239.72959,23623.57194,-10136.544,4455.52859,12460.80809,20394.78994,35415.06254,63689.81811,91951.23817    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 14 | 59,2014,59_2014,0.709703857,1474.870341,17326.9681,15723.4091,29058.81846,16826.96629,-8696.561342,4617.500298,12333.92433,20158.80616,34051.41843,63647.26437,90838.8114    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 15 | 60,2014,60_2014,0.65979272,1238.697279,18019.17038,14654.20243,27343.74841,16549.94393,-10565.82557,4719.201654,12349.03159,20104.97294,33387.03545,63699.47867,92001.931    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 16 | 61,2014,61_2014,0.555983616,1600.024253,16343.18905,14827.64651,23479.28784,15687.4784,-9071.991841,4724.881681,12305.3518,20103.01456,33642.1932,64024.90993,90804.17379    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 17 | 62,2014,62_2014,0.643052412,1176.728447,16764.7812,12941.56662,22222.2386,14844.87791,-8062.887151,4866.156191,12144.61619,20090.66363,33371.35759,64376.86696,90901.6243    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 18 | 63,2014,63_2014,0.516921804,2554.007972,18081.97822,16588.70104,28666.45495,17596.4857,-10794.64224,4711.251963,12395.16856,20131.91147,34092.27103,63988.52421,91312.969    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 19 | 66,2014,66_2014,0.754273459,1220.20486,18386.57227,15563.94269,33125.21131,17610.48131,-8476.443233,4760.572695,12331.76363,20238.29863,33807.34891,63887.15748,90658.08,;   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 20 | 67,2014,67_2014,0.746511466,1149.923491,16554.65646,13733.28984,25581.9361,15866.80592,-9733.36049,4917.306277,12355.77485,19940.4898,33546.76799,63522.30742,90686.3436,;   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 21 | 71,2014,71_2014,0.639647028,1197.454905,14405.61001,13997.83766,25663.40476,14275.83462,-10812.72973,4440.425299,12340.93035,19986.31099,33324.50039,63688.80557,91980.78    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 22 | 78,2014,78_2014,0.637995305,1001.062274,14283.19551,13498.39873,21294.0115,14009.45454,-9412.713216,4787.723227,12241.89117,19831.18583,33495.86923,64217.79214,90341.971    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 23 | 80,2014,80_2014,0.63929261,1280.595362,14571.78418,14380.54187,21844.91401,14728.68475,-10618.16286,5052.528367,12310.11029,19936.00396,33307.18671,64361.1352,89378.0594    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 24 | 87,2014,87_2014,0.560188025,1527.682722,17125.41022,15175.97678,27362.90704,16328.5886,-11650.54054,4678.41356,12356.70458,20019.01995,33975.29075,64189.69932,90556.0914    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 25 | 88,2014,88_2014,0.67559638,1076.311009,13933.30322,13735.50547,26107.95468,13789.43832,-13418.76861,4709.434268,12213.4564,19849.39318,33661.3638,63951.81586,91402.57887    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 26 | 89,2014,89_2014,0.59422618,1084.662254,17192.34833,15606.85164,26649.81217,16260.99878,-10398.18548,4445.654446,12326.67728,20123.24337,34210.69162,63540.01533,90375.394    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 27 | 90,2014,90_2014,0.45553852,1249.750956,17714.85445,16160.90096,27779.15666,17272.50874,-9083.395509,4690.735823,12411.06145,20134.08552,33577.22531,63876.78132,90194.358    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 28 | 100,2014,100_2014,0.889325824,2217.120011,18081.87076,16097.87634,34222.07934,18927.03964,-16830.74334,4869.858079,12548.0876,20099.16735,34363.69451,63575.29873,92031.6    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 29 | 110,2014,110_2014,0.591549039,1684.592153,14441.28799,13997.47063,26702.15848,14217.61712,-12390.84867,4339.125917,12447.22837,19736.23614,33785.61185,63740.47809,91599.    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 30 | 1,2015,1_2015,0.706510266,1592.536365,23240.60107,18768.30811,41848.09605,22383.76337,-10042.25889,4829.327047,12531.63033,20361.95428,34674.82071,63489.98238,91653.8014    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 31 | 7,2015,7_2015,0.757985315,1916.231596,20768.6274,17116.5575,35940.99057,20950.24557,-11534.57567,4677.889928,12581.89411,20360.24955,34450.49282,63612.32101,91367.42961,;   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 32 | 15,2015,15_2015,0.724158814,1686.264274,27872.54539,20669.48635,52227.88874,27165.90551,-10596.05948,4852.080787,12547.3891,20418.84015,35481.51583,63625.35769,91944.717    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 33 | 25,2015,25_2015,0.777355143,1259.788038,21079.42756,15608.84737,38855.9368,19936.81212,-10777.50273,4745.01523,12568.8671,20362.43257,33543.49901,63460.49357,91885.77711    |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |





## LE SURVEY

In ambito delle scienze sociali, il termine **survey** si riferisce a un metodo di ricerca che viene utilizzato per raccogliere dati empirici da un campione rappresentativo di individui o gruppi al fine di ottenere informazioni su atteggiamenti, opinioni, comportamenti o caratteristiche demografiche della popolazione di riferimento. Questo metodo è ampiamente utilizzato nella sociologia, nella psicologia, nell'economia, nella scienza politica e in molte altre discipline delle scienze sociali.

Un survey di solito coinvolge la somministrazione di un questionario o un'intervista a un gruppo selezionato di partecipanti, chiamato campione. Il campione dovrebbe essere scelto in modo da essere rappresentativo della popolazione di interesse, al fine di estrarre conclusioni generali o fare previsioni sul comportamento o le opinioni dell'intera popolazione.

I sondaggi possono essere condotti in vari formati, tra cui sondaggi telefonici, questionari online, interviste faccia a faccia o questionari cartacei. L'analisi dei dati raccolti dai sondaggi può includere statistiche descrittive, analisi delle tendenze e, in alcuni casi, l'utilizzo di metodi statistici avanzati per trarre conclusioni significative.

L'obiettivo principale di un survey nelle scienze sociali è quello di raccogliere dati affidabili e validi che possano essere utilizzati per comprendere meglio il comportamento umano, le opinioni, le tendenze sociali e altre variabili rilevanti all'interno di una determinata popolazione. I risultati dei sondaggi possono essere utilizzati per prendere decisioni politiche, pianificare strategie di marketing, condurre ricerche accademiche o rispondere a domande di interesse sociale.

## LE MAPPE 1/2

### Definizione generale delle "mappe" nelle scienze sociali:

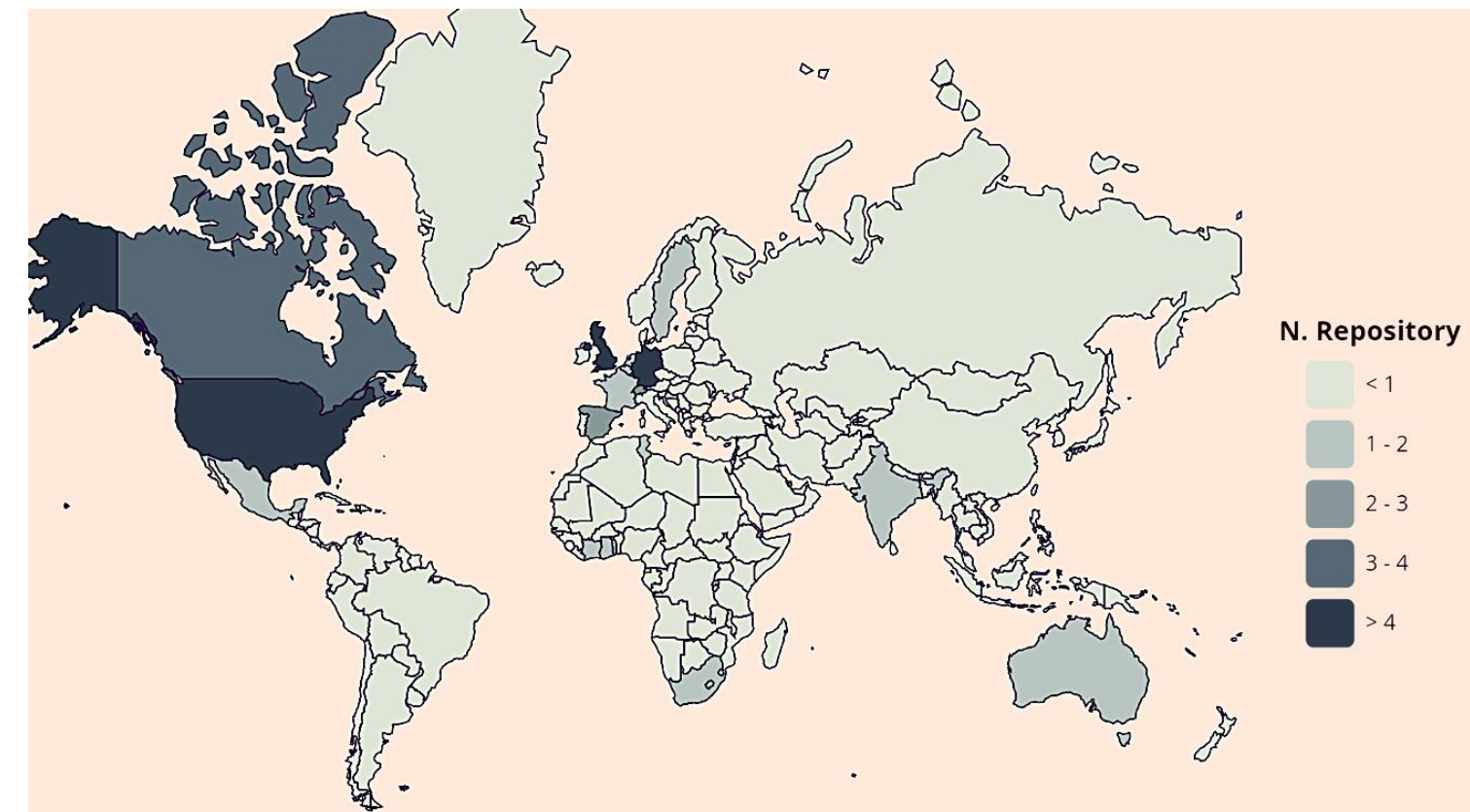
Le mappe nelle scienze sociali si riferiscono a rappresentazioni visive o grafiche di dati, concetti o relazioni tra variabili sociali. Queste mappe sono utilizzate per esplorare, analizzare e comunicare informazioni complesse relative a fenomeni sociali, politici o economici. Possono essere strumenti utili per rappresentare dati geospaziali, reti sociali, flussi migratori, relazioni culturali o qualsiasi altro aspetto studiato nelle scienze sociali.

### Formati comuni di input per la creazione di mappe nelle scienze sociali:

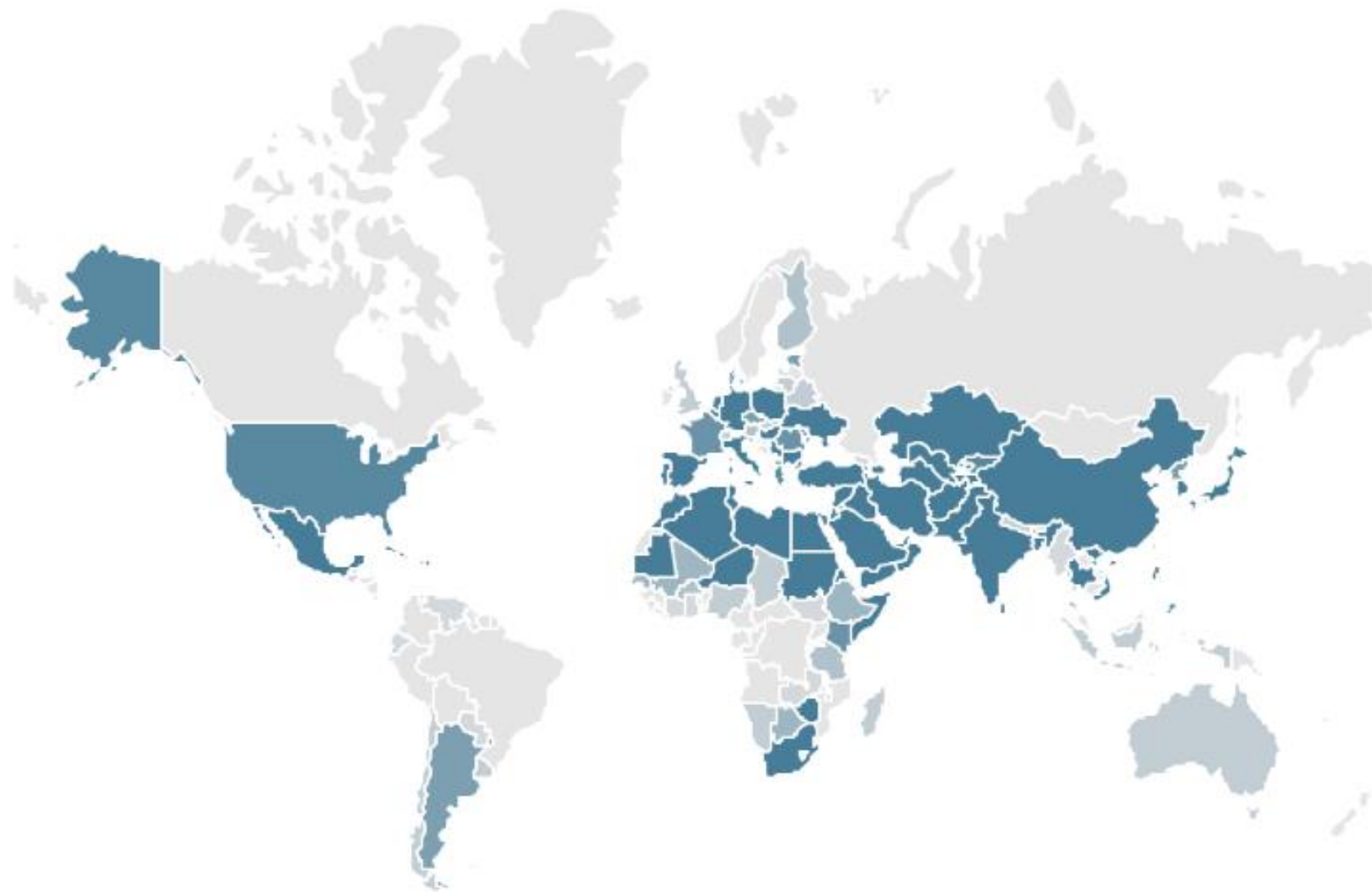
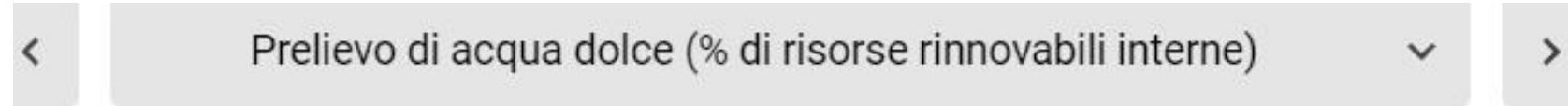
**Dati tabellari:** Molto spesso, i dati sociali vengono organizzati in tabelle, ad esempio in fogli di calcolo Excel o database. Questi dati possono includere informazioni geografiche (ad esempio, dati demografici per regioni), dati di rete sociale (ad esempio, relazioni tra individui) o qualsiasi altra variabile sociale di interesse.

**Dati geospaziali:** Per la creazione di mappe geospaziali, vengono utilizzati dati geografici, come coordinate geografiche (latitudine e longitudine), dati cartografici o confini amministrativi. Questi dati possono essere ottenuti da fonti pubbliche o raccolti attraverso rilevamenti sul campo.

**Dati di rete:** Quando si creano mappe delle reti sociali o delle relazioni tra attori sociali, i dati di input spesso includono informazioni sulle connessioni tra gli individui o le entità. Questi dati sono generalmente organizzati in forma di grafi.



## LE MAPPE 2/2



### Formati comuni di output per la creazione di mappe nelle scienze sociali:

**Mappe cartografiche:** Queste sono rappresentazioni visive delle informazioni geospaziali. I formati comuni includono mappe cartacee, mappe digitali interattive (come quelle create con software GIS, Geographic Information Systems), o formati web come GeoJSON.

**Grafi sociali:** Quando si rappresentano reti sociali, i formati di output possono includere grafi, diagrammi a dispersione o rappresentazioni visive delle connessioni tra gli attori sociali.

**Diagrammi e grafici:** Spesso, i risultati delle analisi sociali vengono rappresentati sotto forma di diagrammi a barre, grafici a dispersione, grafici a torta o altre rappresentazioni grafiche per comunicare i risultati in modo efficace.

**Software di visualizzazione dei dati:** L'output può essere generato attraverso software di visualizzazione dei dati, come Tableau, Power BI o D3.js, che consentono di creare visualizzazioni interattive personalizzate.

In generale, la scelta del formato di input e output per la creazione di mappe nelle scienze sociali dipende dalla natura dei dati e dagli obiettivi della ricerca. L'obiettivo principale è creare rappresentazioni visive chiare e significative dei dati sociali per facilitare la comprensione e la comunicazione delle informazioni.





## GLI ALGORITMI

Gli **algoritmi** nel contesto della ricerca nelle scienze sociali sono procedure o istruzioni sistematiche che guidano il processo di analisi e interpretazione dei dati sociali. Questi algoritmi sono progettati per estrarre informazioni significative dai dati, identificare pattern, tendenze o relazioni, e supportare la ricerca scientifica nell'ambito delle scienze sociali. Ecco cinque esempi di software che utilizzano algoritmi nella ricerca sociale:

### **SPSS (Statistical Package for the Social Sciences):**

SPSS è un software ampiamente utilizzato nelle scienze sociali per l'analisi statistica dei dati. Offre una vasta gamma di algoritmi statistici, tra cui regressione, analisi di varianza, analisi fattoriale e molti altri, che consentono agli studiosi di eseguire analisi complesse sui dati sociali.

### **STATA:**

STATA è un altro popolare software statistico utilizzato nella ricerca sociale. Fornisce una serie di algoritmi per l'analisi dei dati longitudinali, l'analisi multinomiale, la regressione robusta e molte altre tecniche statistiche.

### **R:**

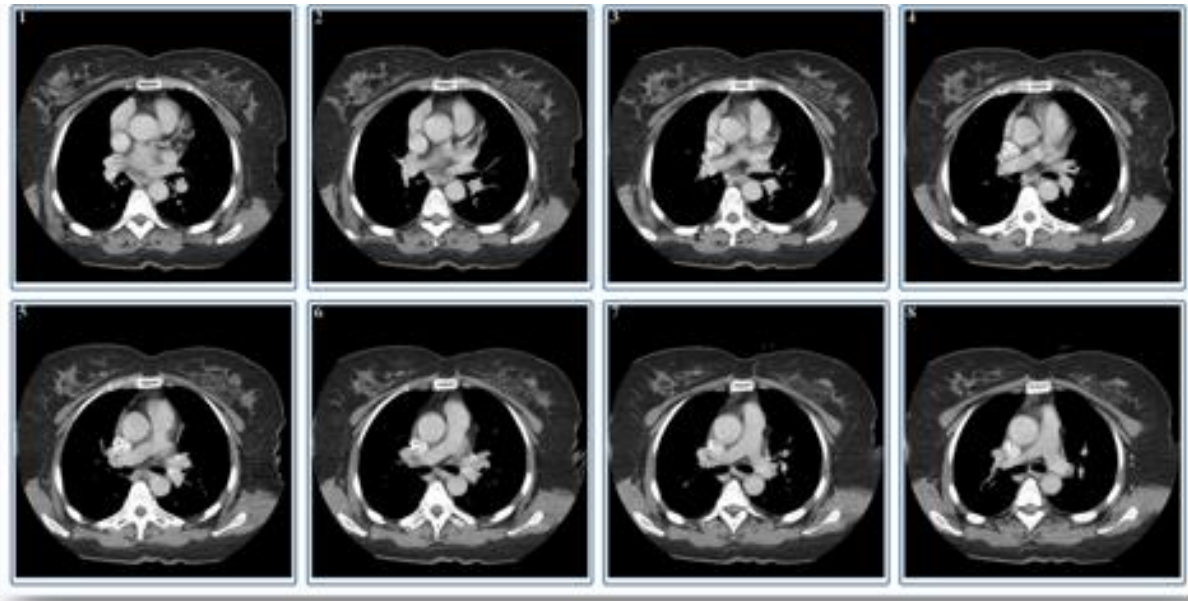
R è un linguaggio di programmazione e un ambiente di sviluppo statistico open-source. È estremamente flessibile e offre una vasta gamma di pacchetti e librerie che contengono algoritmi per l'analisi dei dati sociali, tra cui regressione, analisi dei cluster, analisi di rete e altro ancora.

### **Python:**

Python è un altro linguaggio di programmazione popolare che è ampiamente utilizzato nella ricerca sociale

## QUALI TIPI DI DATI...

... dipende dal dominio di ricerca.



Per un ricercatore di area biomedica il dato può essere l'immagine di una TAC



Per un ricercatore di ingegneria il dato può essere un algoritmo

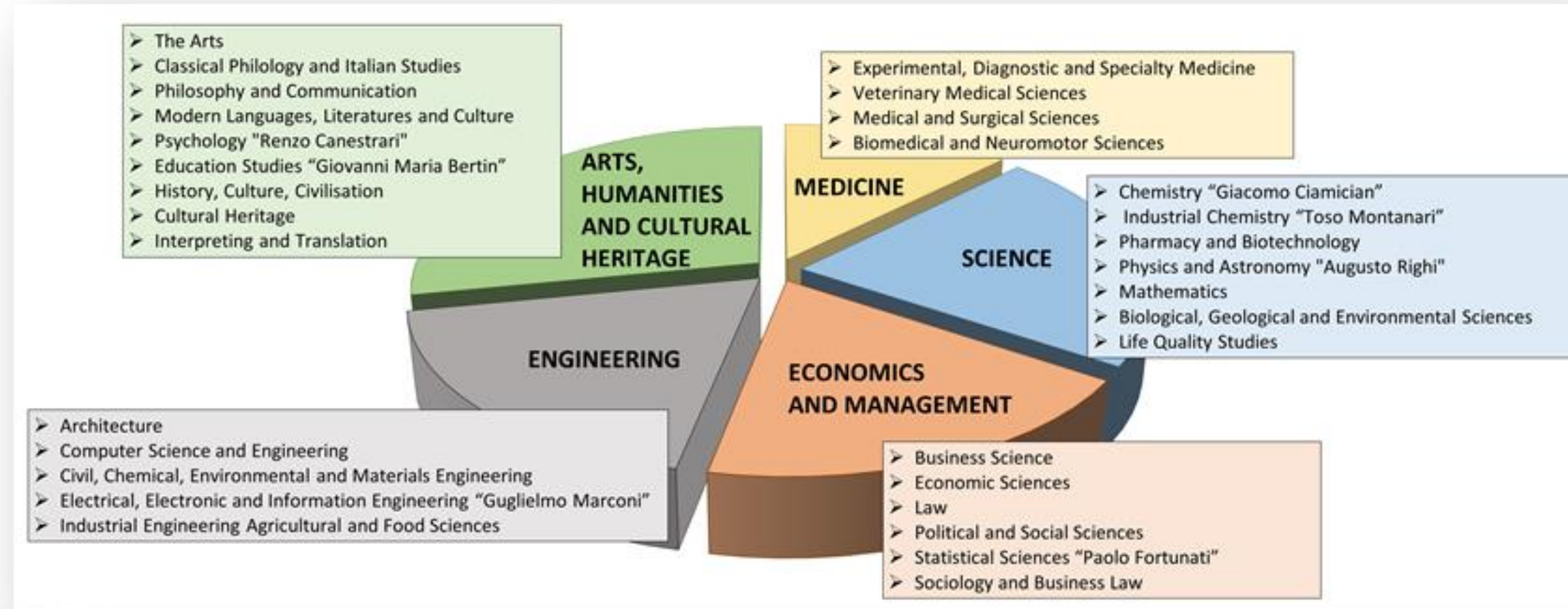


Per un ricercatore di area umanistica il dato può essere una fonte testuale antica

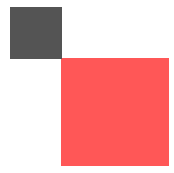
## DIVERSI TIPI DI DATI...

... si incontrano in un contesto sempre più multidisciplinare.

Come quello della nostra Università, che include 32 dipartimenti raggruppati in 5 macroaree.



Anche a livello di enti finanziatori, è incentivata la collaborazione tra ricercatori anche di diversi ambiti.



# GESTIRE IL DATO DELLA RICERCA







La gestione dei dati di ricerca (Research Data Management, RDM) è la gestione e l'organizzazione attenta dei dati di ricerca durante l'intero ciclo di ricerca, con l'obiettivo di rendere il processo di ricerca il più efficiente possibile e di facilitare la cooperazione con gli altri.



## I VANTAGGI DI UNA CORRETTA GESTIONE DEL DATO

I dati che non sono gestiti hanno un'appetibilità bassa per lo stesso ricercatore che li produce.



## DATA MANAGEMENT HORROR STORIES

---

"Ho comprato un portatile nuovo di zecca solo per le mie ricerche (...).  
Lo usavo per caricare i dati della ricerca (interviste e diari sul campo) e  
tutto ciò che riguardava la ricerca (...).  
Ho dimenticato il portatile da qualche parte (...).

I dati erano sensibili e non erano memorizzati altrove (...).

Il collega ricercatore ha comunque osato chiedermi se avessi fatto un  
Data Management Plan e una valutazione dei rischi. Mi avrebbe  
aiutato ad anticipare i problemi."

<https://research.tuni.fi/researchdata/news/data-horror-week-end-of-stories/>



## DATA MANAGEMENT HORROR STORIES

---



"Abbiamo raccolto dati di ricerca ricchi e diversificati(...). Poi il progetto è terminato e ci è venuto in mente che potevamo depositare i dati per un ulteriore utilizzo (...).

Hanno iniziato a farci domande difficili come: avete chiesto ai partecipanti alla ricerca il permesso di archiviare i dati e avete menzionato l'archiviazione nella nota sulla privacy? (...) Ci era sfuggito tutto questo perché avevamo fretta di iniziare il lavoro sul campo. Di conseguenza, abbiamo perso l'opportunità di aprire i dati ad altri e persino di utilizzarli nei nostri progetti futuri."

<https://research.tuni.fi/researchdata/news/data-horror-week-end-of-stories/>



## DATA MANAGEMENT HORROR STORIES

---



"Abbiamo diversi ricercatori in questo stesso progetto e stiamo tutti lavorando sugli stessi dati. Qualcuno sfortunato (non io) ha toccato il file dei dati grezzi e vi ha apportato alcune modifiche. Non ci sono più i dati originali. Allo stesso modo, non è chiaro chi abbia fatto cosa e quando con i dati.

Come potremo ora ottenere una descrizione valida dell'elaborazione dei dati in una rivista scientifica e vogliamo che questo studio sia trasparente? Gente, tenete i dati grezzi separati e le regole del gioco chiare."

<https://research.tuni.fi/researchdata/news/data-horror-week-end-of-stories/>

## DATA MANAGEMENT HORROR STORIES

---

"Ho sviluppato quella che pensavo fosse una struttura logica di cartelle, c'erano sottocartelle di sottocartelle e le loro sottocartelle, e le cartelle avevano nomi con abbreviazioni fantasiose (...). Poi ho dovuto prendermi una pausa dalla ricerca (...).

È stato un po' angosciante tornare lì e riprendere in mano i miei dati. Beh, cartella dopo cartella e file dopo file tutto ha cominciato a funzionare."

<https://research.tuni.fi/researchdata/news/data-horror-week-end-of-stories/>



---

## QUALI VANTAGGI PUOI AVERE CON UNA CORRETTA GESTIONE DEL DATO DI RICERCA?



Organizzare i dati rende il tuo lavoro più efficiente.

In termini di costi/tempo: i dati gestiti una volta restano interpretabili, comprensibili e rintracciabili.



Se li gestisci, potresti non perderli.

L'archiviazione corretta dei dati e il backup regolare prevengono le perdite di dati.



Alcuni dati potrebbero essere unici e non riproducibili.

Questo li rende preziosissimi per la comunità scientifica.



Aumenta l'integrità della ricerca.

Un dato correttamente gestito facilita la validazione e il controllo.



Stimola la collaborazione con altri ricercatori.

Troveranno più facile comprendere e riutilizzare i vostri dati.



## Per i ricercatori

- Maggiore visibilità e citazioni
- Opportunità di collaborazione
- Riconoscimento per la carriera
- Aiuta a prevenire errori e migliora la qualità delle analisi dei dati
- Diminuisce i rischi di "non-compliance" (etica, privacy, policy degli enti finanziatori)



## Per la ricerca

- Facilita la rintracciabilità e il riuso dei dati
- Consente nuove ricerche e nuovi approfondimenti sui dati
- Protezione dei dati di valore
- Sostiene l'integrità e la riproducibilità della ricerca

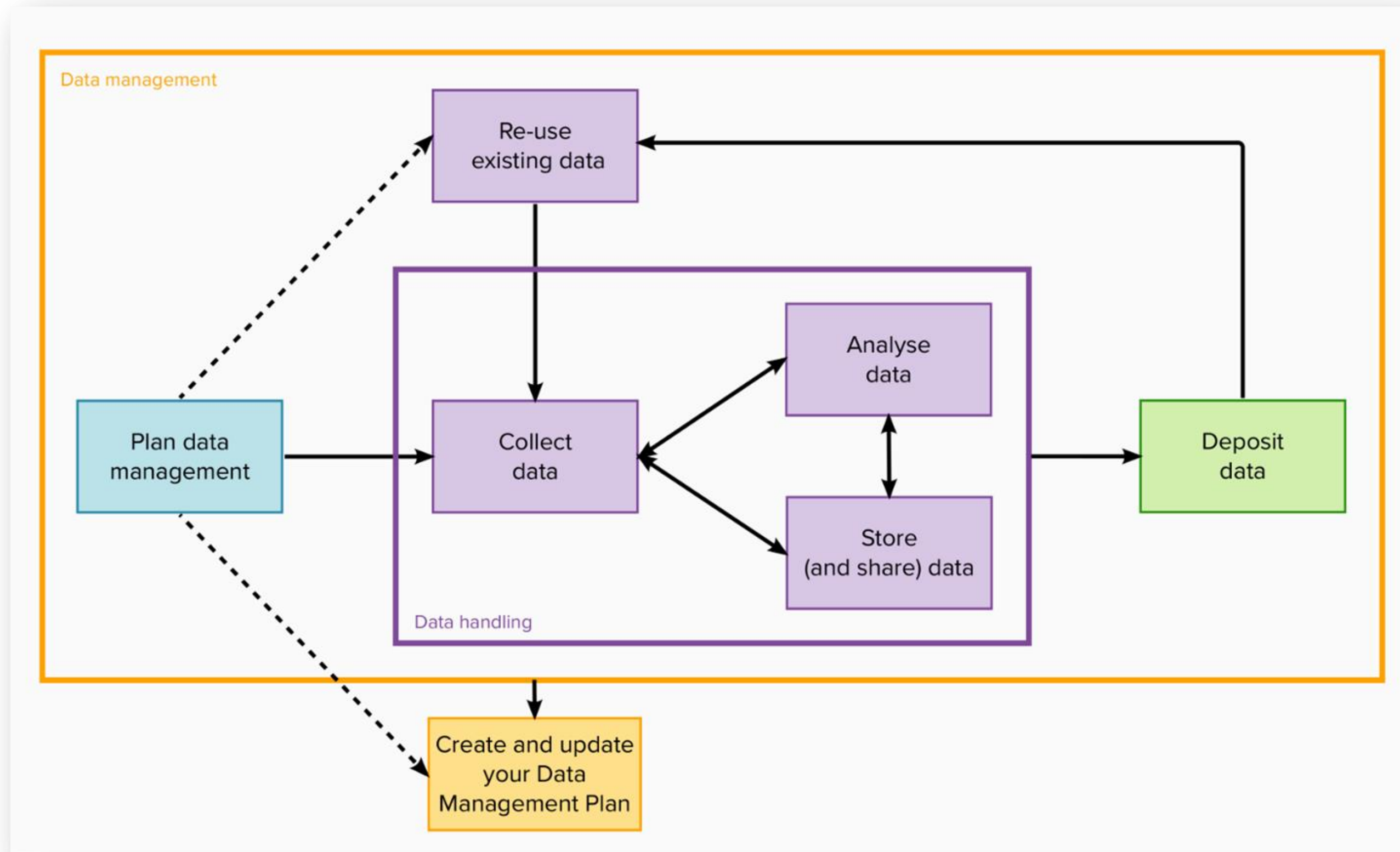


## Per la società

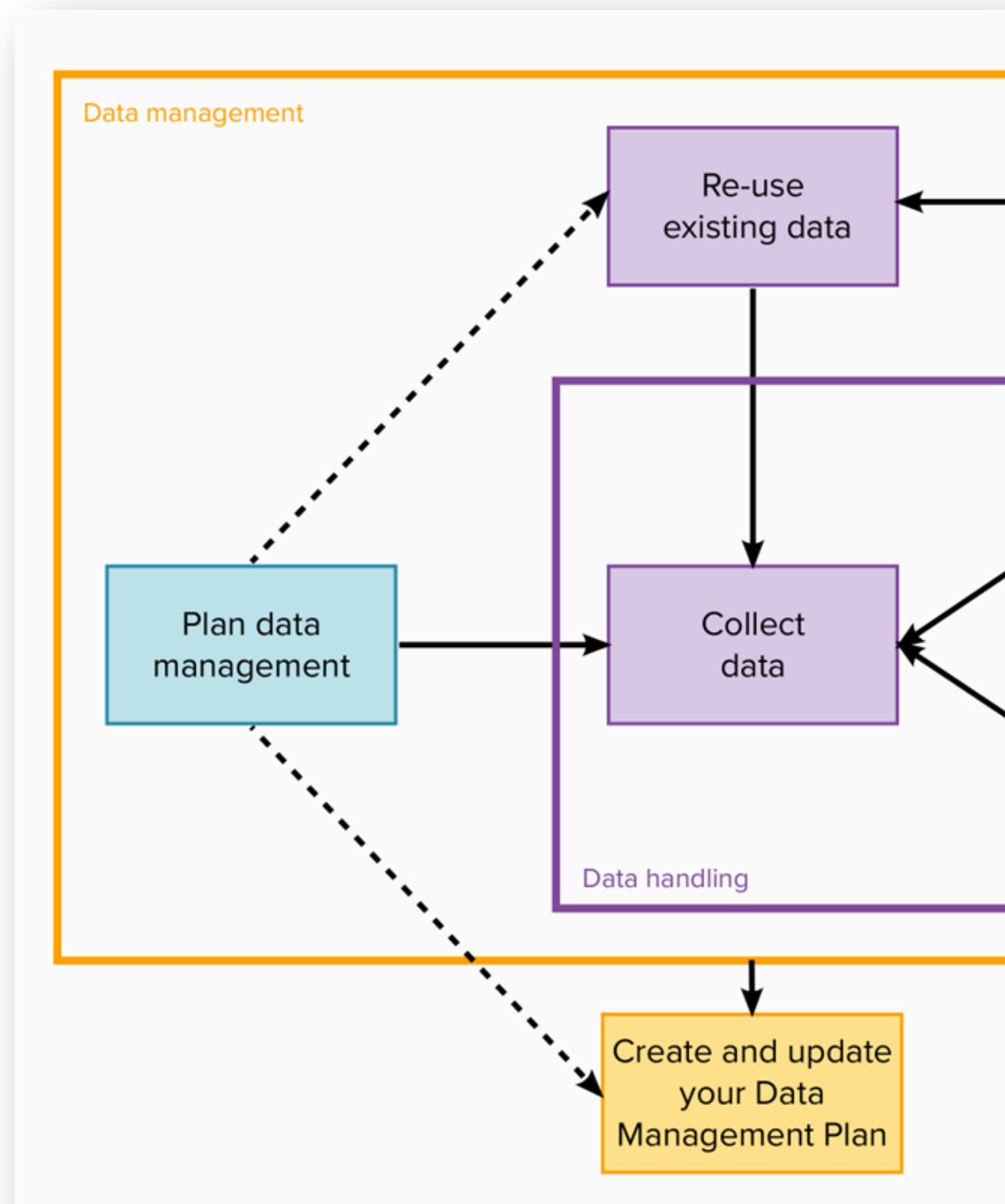
- Uso efficiente delle risorse pubbliche
- Una migliore qualità della ricerca può dare beneficio al processo decisionale
- Opportunità di business
- Opportunità di citizen science
- Maggiore trasparenza e fiducia nella scienza



## LE FASI DELLA GESTIONE DEL DATO DI RICERCA: IL CICLO DI VITA DEL DATO



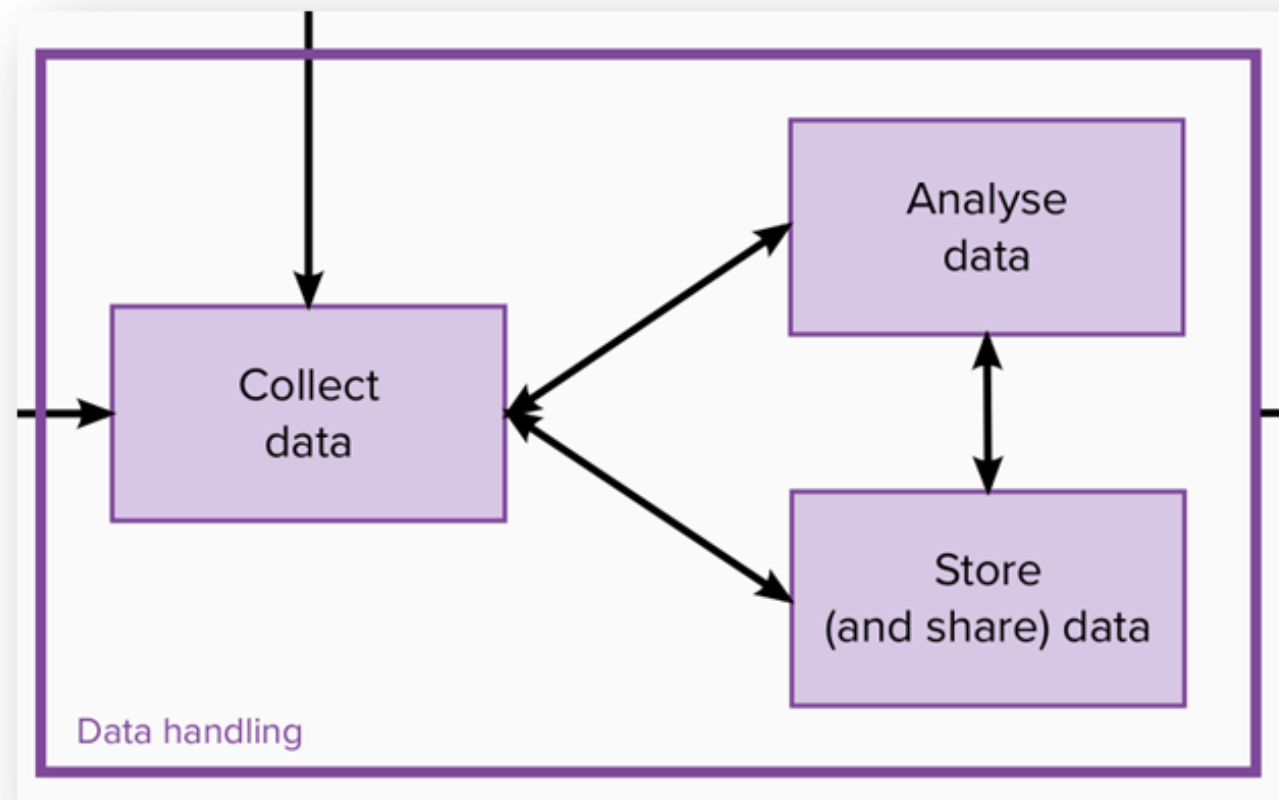
## LE FASI DELLA GESTIONE DEL DATO DI RICERCA: IL CICLO DI VITA DEL DATO



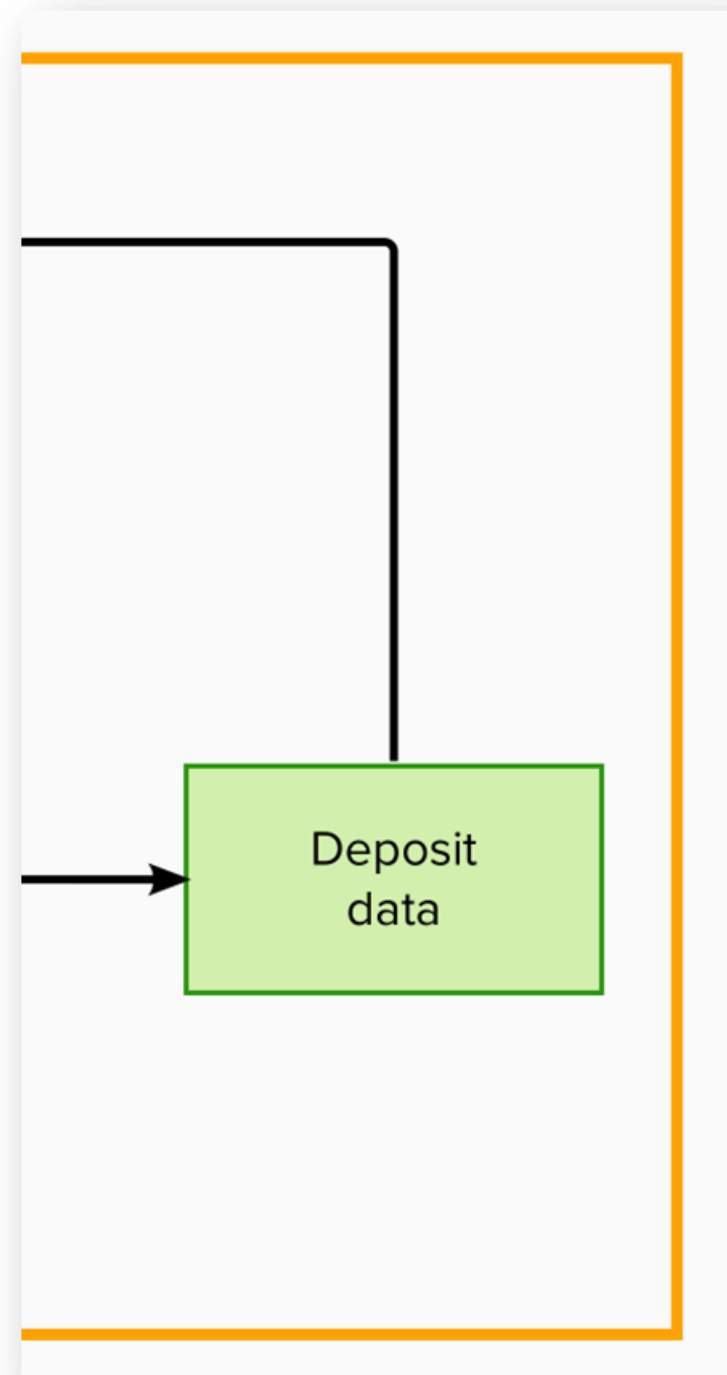
- La fase di pianificazione è cruciale per operare delle scelte ragionate durante tutto lo svolgimento del progetto.
- I dati possono essere sia generati *ex-novo* durante lo svolgimento del progetto che riutilizzati da fonti già esistenti.
- La stesura di un Data Management Plan è una buona pratica per tenere traccia dall'inizio del Progetto delle scelte fatte in termini di gestione del dato.

## LE FASI DELLA GESTIONE DEL DATO DI RICERCA: IL CICLO DI VITA DEL DATO

- Le fasi attive della ricerca sono le più concitate, in cui è più facile perdere di vista la corretta gestione del dato per arrivare velocemente ad un risultato.
- Portare avanti processi di controllo qualità dei dati e tenere traccia delle diverse versioni è fondamentale per evitare errori e lavorare sempre con dati consistenti.
- In questa fase si possono scegliere i formati più adatti ai propri dati.
- Durante le fasi attive della ricerca i dati devono essere conservati in spazi di archiviazione appropriati, provvedendo ai backup necessari per non rischiare di perdere i propri dati.
- Organizzare file e cartelle in modo gerarchico e assegnando nomi coerenti aiuta a non perdersi nella navigazione!



## LE FASI DELLA GESTIONE DEL DATO DI RICERCA: IL CICLO DI VITA DEL DATO



- Il deposito dei dati, organizzati in datasets, è fondamentale per la preservazione a lungo termine e per la condivisione dei dati della ricerca.
- Depositare i dati nei repositories adeguati ne garantisce la visibilità da parte di altri ricercatori, favorisce la riproducibilità e aumenta il valore della ricerca.

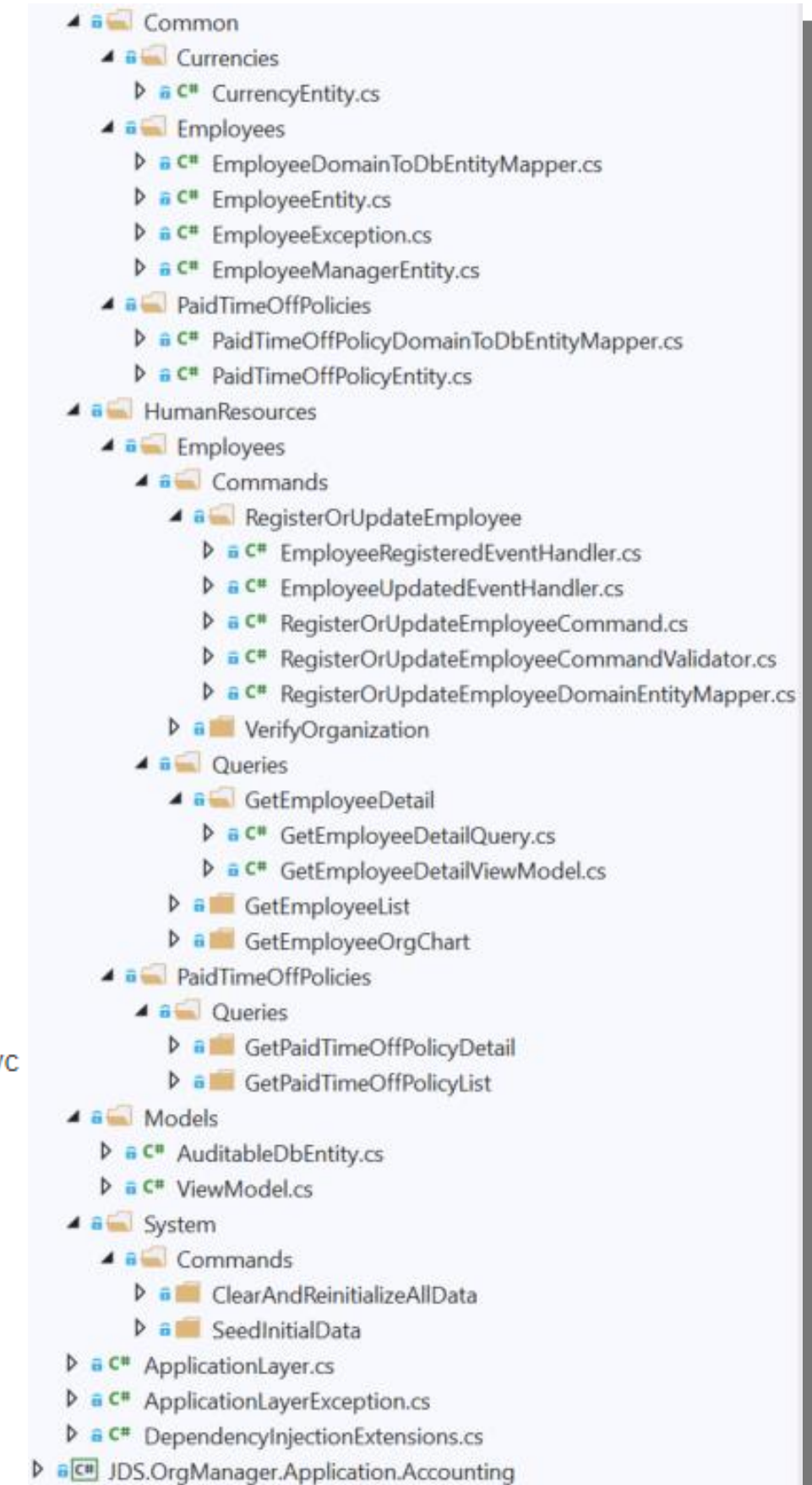
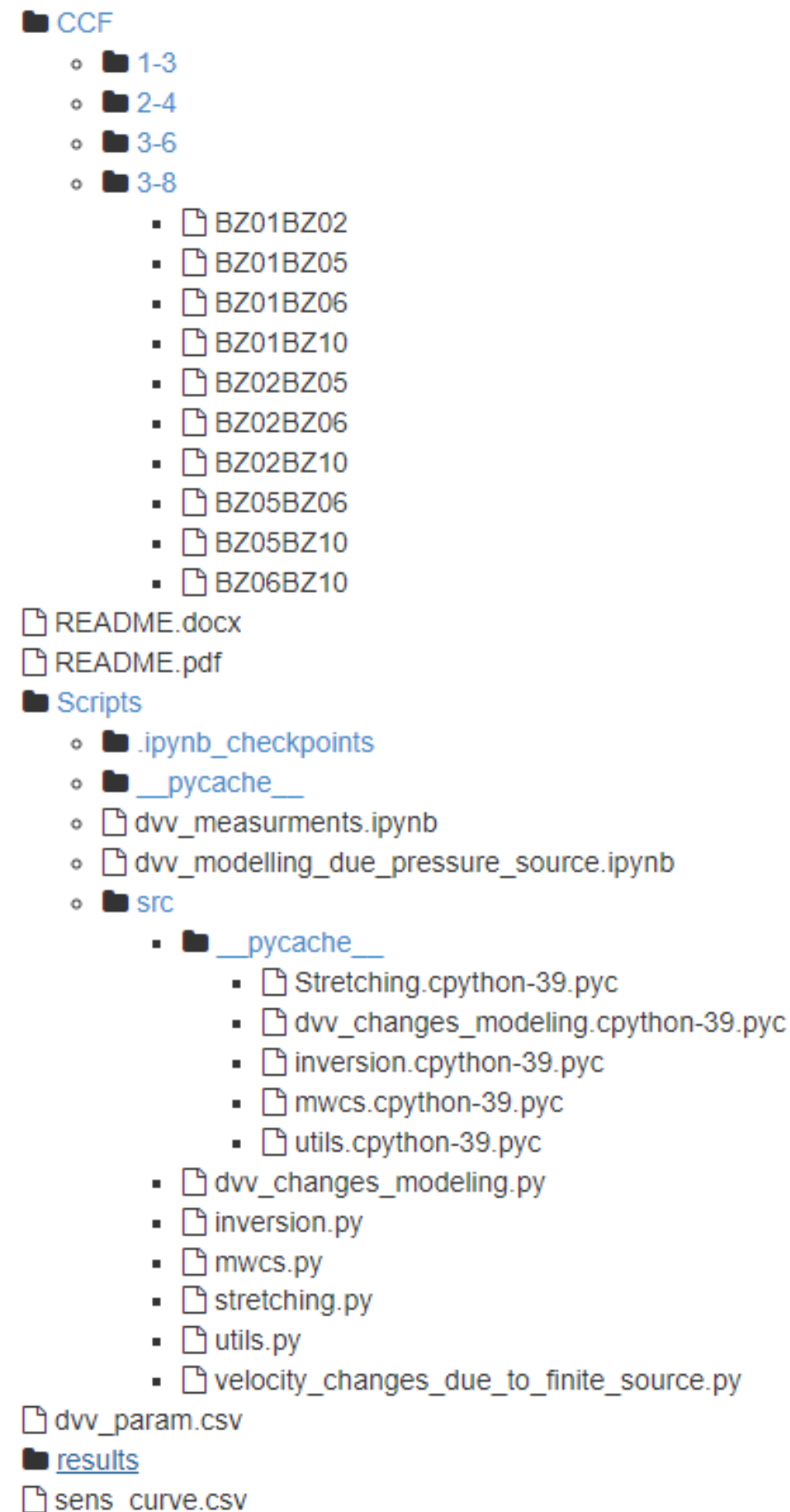
NB. Con «deposito dei dati» non si intende il loro caricamento su siti web personali, siti web degli editori, servizi di cloud storage (Dropbox, Google drive, ecc.), piattaforme di ricerca (Academia.edu, ResearchGate)!



## I DATASET

DATI: informazioni raccolte per essere esaminate e utilizzate come base per ragionamenti, discussioni o calcoli e, più in generale, utilizzate per convalidare un'ipotesi.

DATASET: raccolta di dati assemblati con uno scopo comune e organizzati per riflettere i risultati di un'attività di ricerca.





# QUALCHE PUNTO CHIAVE

- Esistono diverse definizioni di dato e variano in base alla disciplina di riferimento.
- Dato è tutto ciò che è alla base dei ragionamenti a supporto di una tesi di ricerca.
- Una collezione di dati accomunati dallo stesso obiettivo è chiamata dataset.
- Il dato ha un valore intrinseco come asset della ricerca, per conservarlo deve essere gestito correttamente.
- La gestione corretta è un processo che attraversa tutte le fasi del ciclo di vita del dato della ricerca.









# RENDERE DISPONIBILI I DATI AUMENTA IL VALORE DELLA RICERCA

## Data and Code Availability Policy

The Econometric Society (ES) and its journals, **Econometrica**, **Quantitative Economics**, and **Theoretical Economics**, have the policy to publish papers that include empirical, experimental, and/or simulation results only if the data and code used in the analysis are clearly and precisely documented and access to the data and code is non-exclusive to the authors.

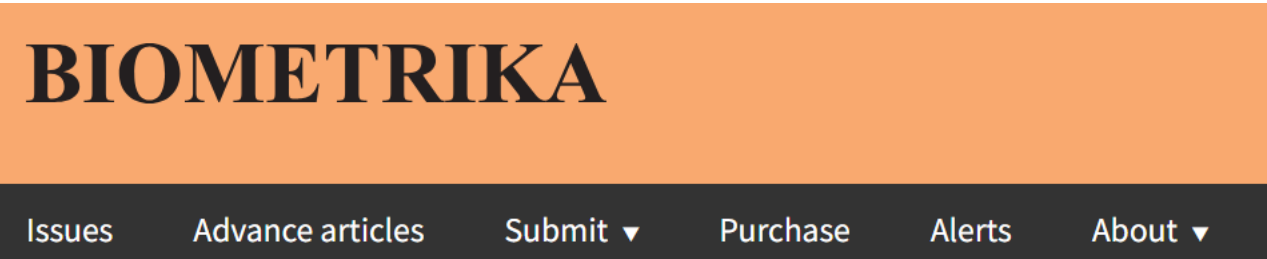
Authors of these papers must provide, prior to acceptance, raw data, codes, and sufficient documentation to permit the replication of all the results in the paper and in the appendices approved for publication online. They must also provide sufficient information to replicate the process of obtaining this raw data from the original sources and cite all the sources of data appropriately.

The journals of the ES will conduct reproducibility checks for the empirical, experimental, and/or simulation results included in the paper and in the approved online appendices prior to final acceptance.

Requests for an exemption from providing the materials described in this policy, or for restricting their usage, should be stated clearly when the paper is first submitted for review. It will be at the editors' discretion whether the paper can then be reviewed. Exceptions will not be considered later in the review and publication process.

By submitting to any journal of the Econometric Society, authors indicate their acceptance of this Data and Code Availability Policy.

<https://www.econometricsociety.org/publications/es-data-editor-website/data-and-code-availability-policy>



## Availability of Data and Materials

Where ethically feasible, Biometrika strongly encourages authors to make all data and software code on which the conclusions of the paper rely available to readers. We suggest that data be presented in the main manuscript or additional supporting files, or deposited in a public repository whenever possible. For information on general repositories for all data types, and a list of recommended repositories by subject area, please see [Choosing where to archive your data](#).

## Data Citation

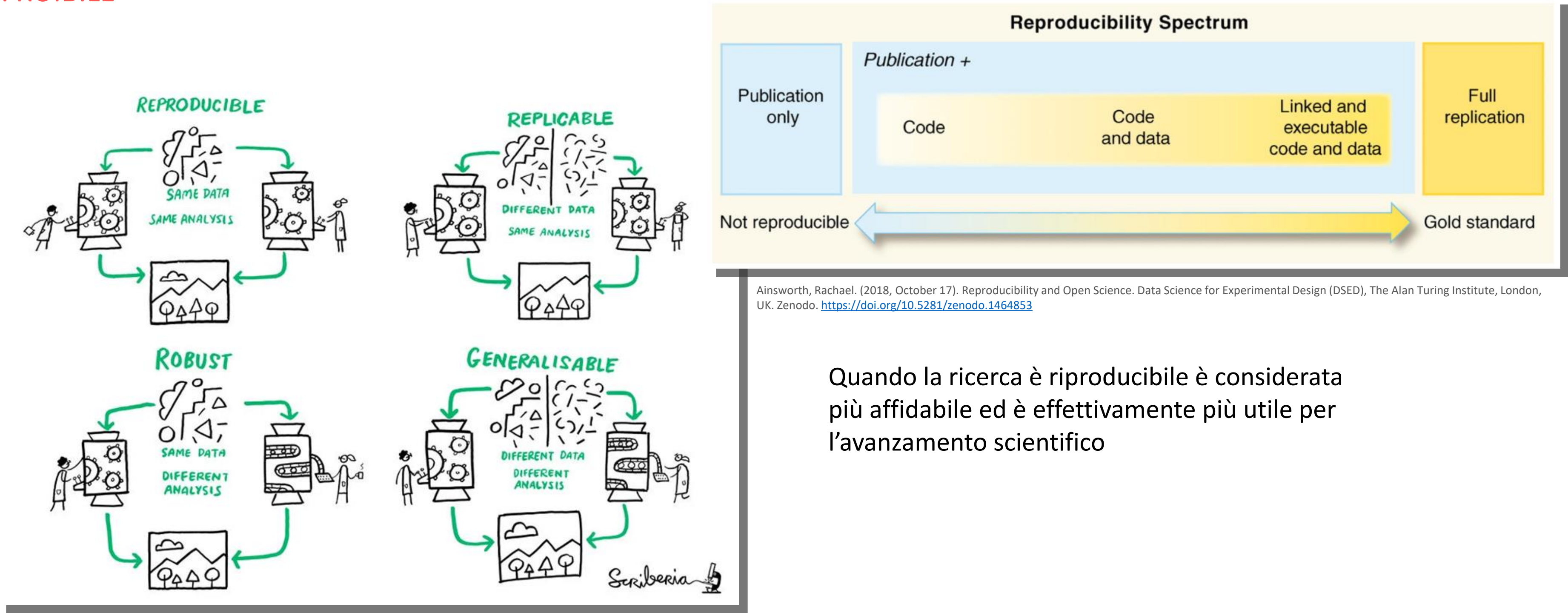
Biometrika supports the [Force 11 Data Citation Principles](#) and requires that all publicly available datasets be fully referenced in the reference list with an accession number or unique identifier such as a digital object identifier (DOI). Data citations should include the minimum information recommended by [DataCite](#):

- [dataset]\* Authors, Year, Title, Publisher (repository or archive name), Identifier

\*The inclusion of the [dataset] tag at the beginning of the citation helps us to correctly identify and tag the citation. This tag will be removed from the citation published in the reference list.

[https://academic.oup.com/biomet/pages/General\\_Instructions](https://academic.oup.com/biomet/pages/General_Instructions)

LA RICERCA RIPRODUCIBILE: AFFIDABILE E FRUIBILE

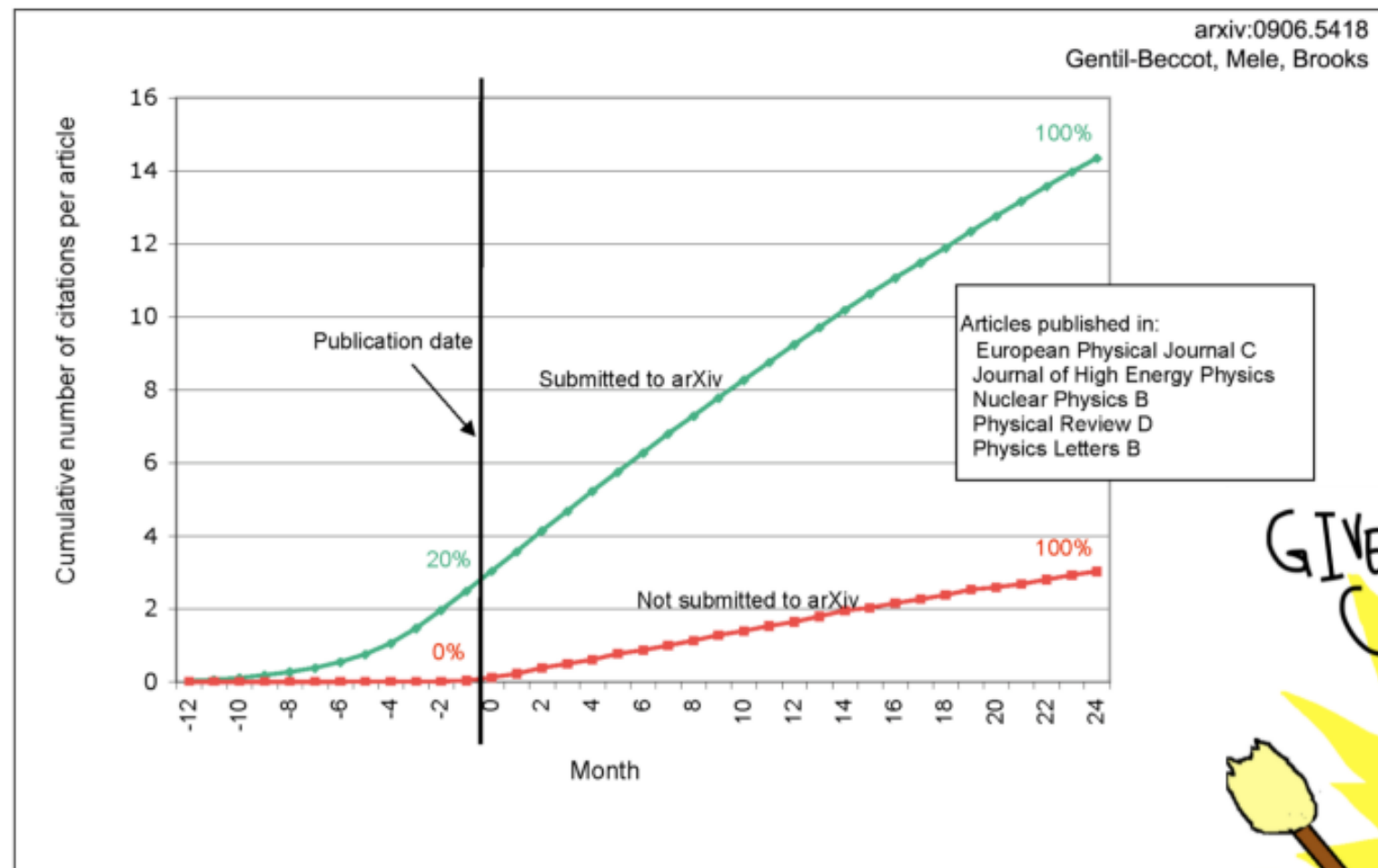


Ainsworth, Rachael. (2018, October 17). Reproducibility and Open Science. Data Science for Experimental Design (DSED), The Alan Turing Institute, London, UK. Zenodo. <https://doi.org/10.5281/zenodo.1464853>

Quando la ricerca è riproducibile è considerata più affidabile ed è effettivamente più utile per l'avanzamento scientifico



## I VANTAGGI PER I RICERCATORI



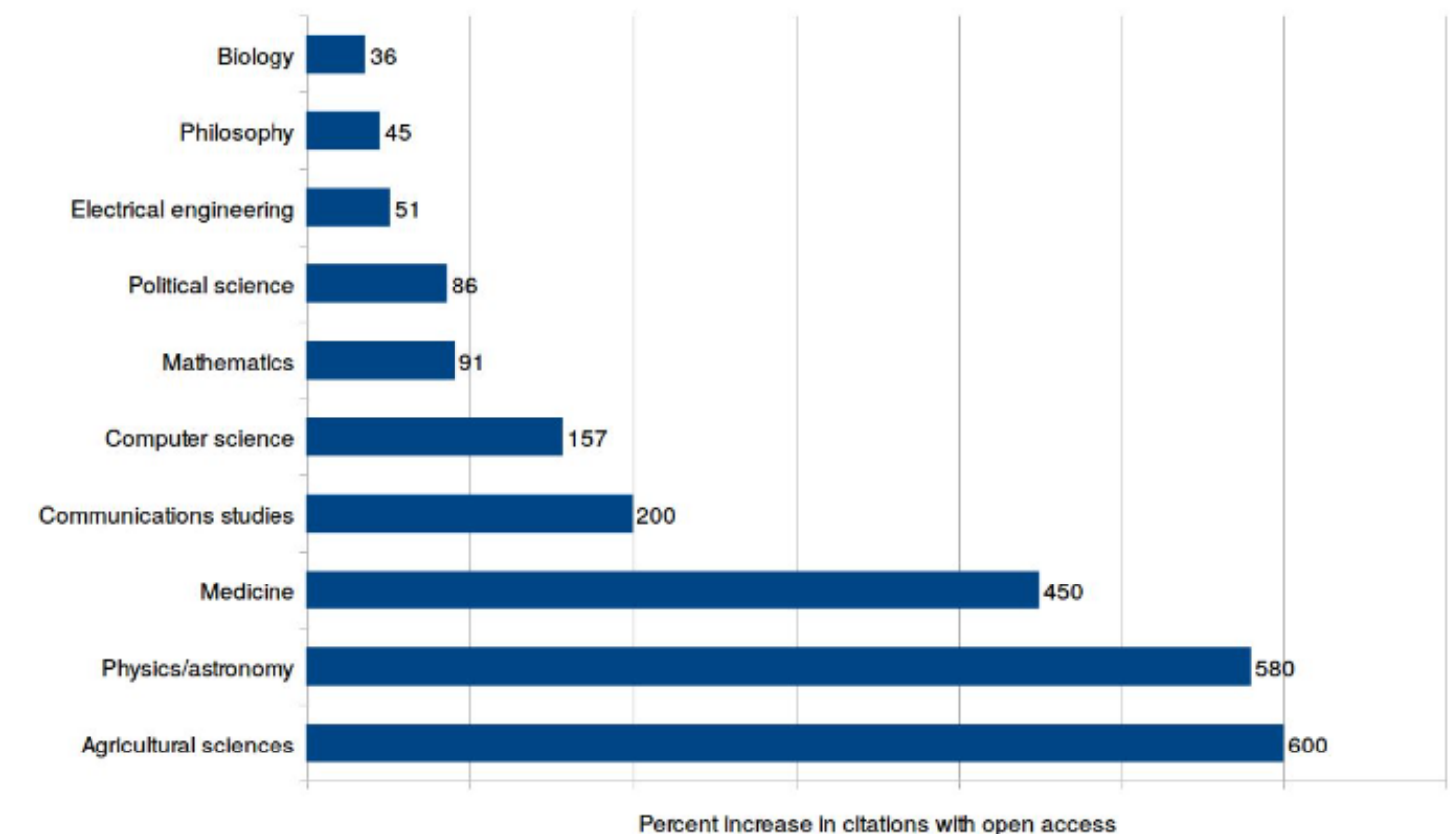
Gentil-Beccot, Mele, Brooks (2009), <https://arxiv.org/abs/0906.5418>

R Ainsworth - @rachaelevelyn



## Why should you share preprints?

- Enable **rapid** publication
- Enable **open access**
- **Citation advantage**

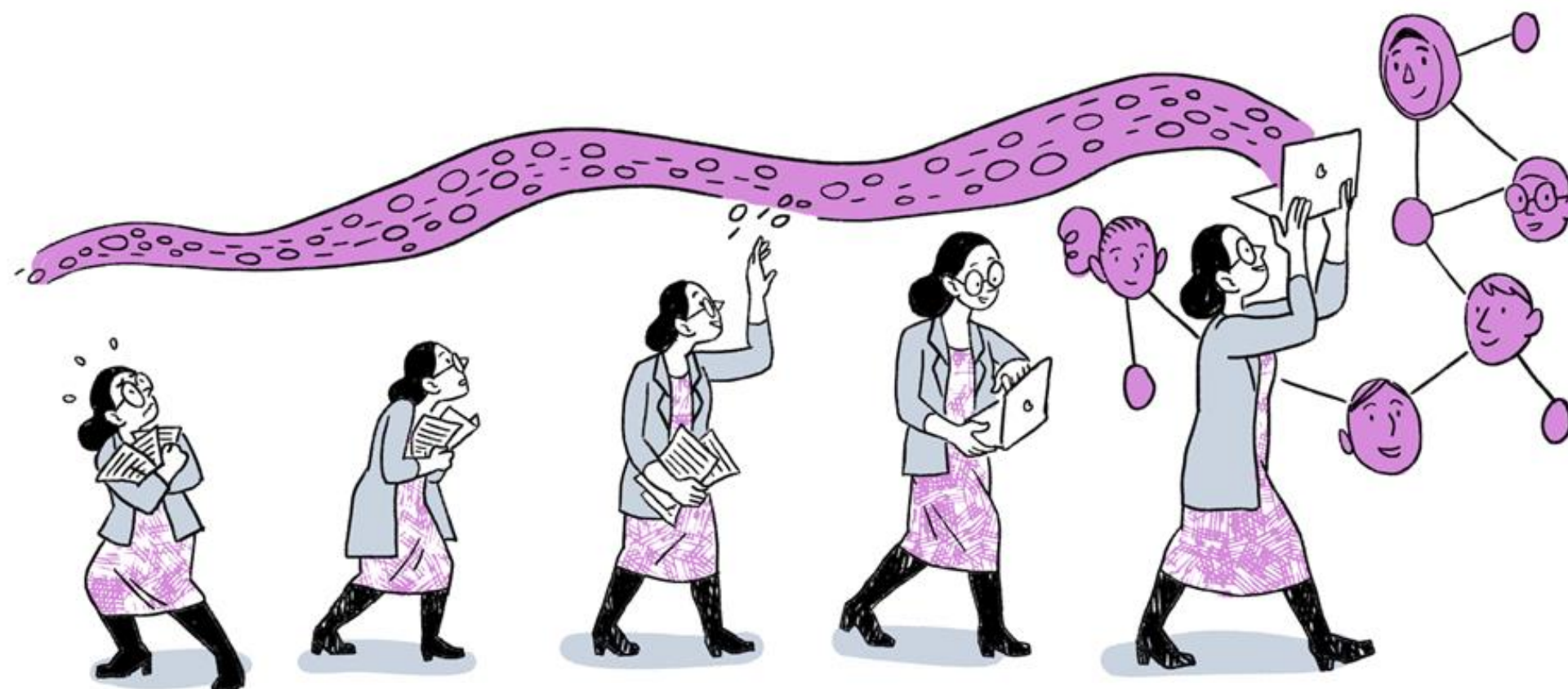


Source: Data from [Alma Swan, 2010](#). Figure produced by E.C. McKiernan (CC BY).

[kevin.moerman@nuigalway.ie](mailto:kevin.moerman@nuigalway.ie)

@KMMoerman

Maggior visibilità e potenziale di collaborazione, networking ed eventualmente finanziamenti



# EVOLVING TOWARDS AN ERA OF OPEN RESEARCH

Scriberia 



COS'È OPEN SCIENCE



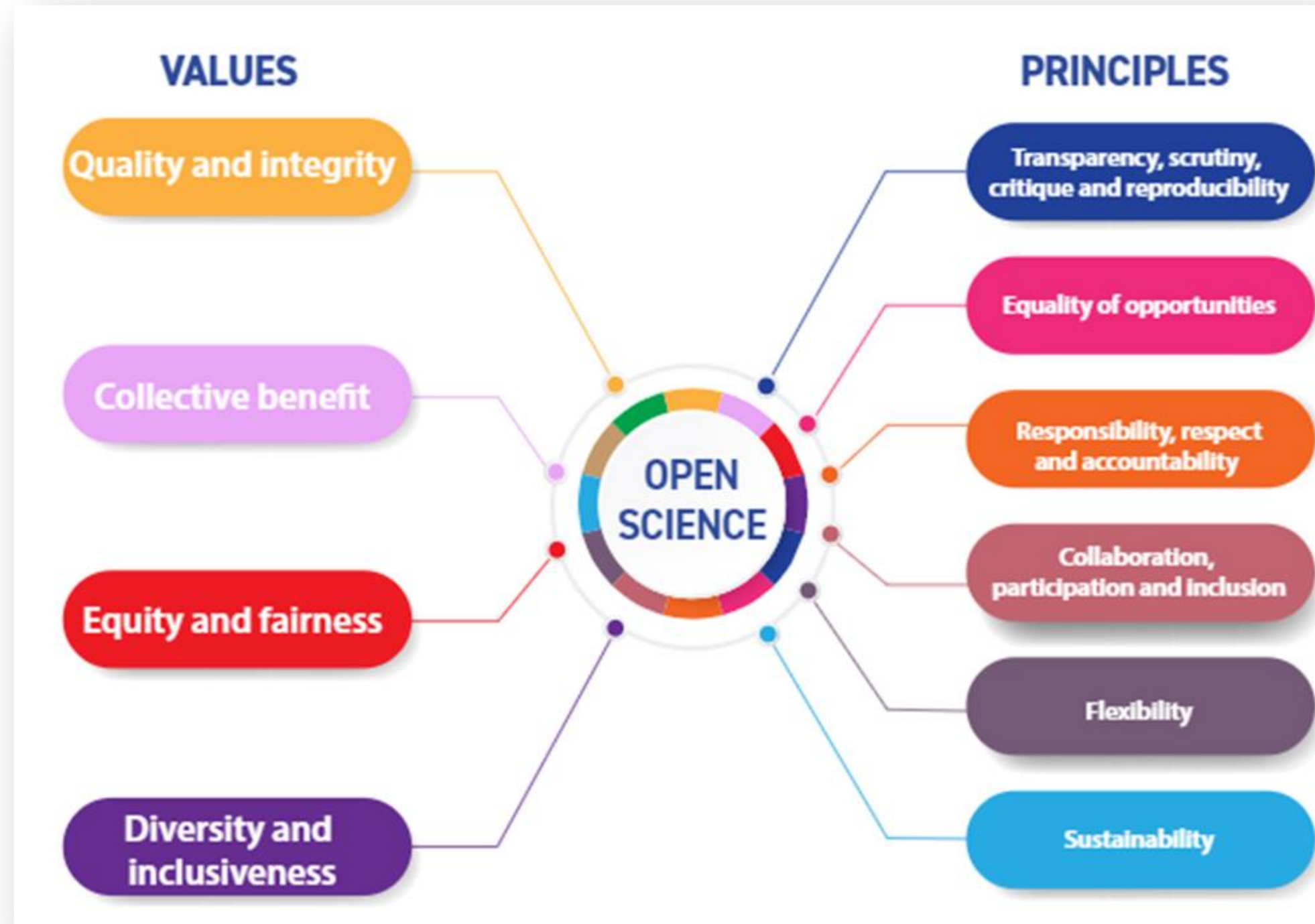


Open Science è definita come un costrutto inclusivo che combina vari movimenti e pratiche che mirano a rendere la conoscenza scientifica multilingue apertamente disponibile, accessibile e riutilizzabile per tutti, ad aumentare le collaborazioni scientifiche e la condivisione delle informazioni a beneficio della scienza e della società e ad aprire i processi di creazione, valutazione e comunicazione della conoscenza scientifica agli attori della società al di là della comunità scientifica tradizionale.





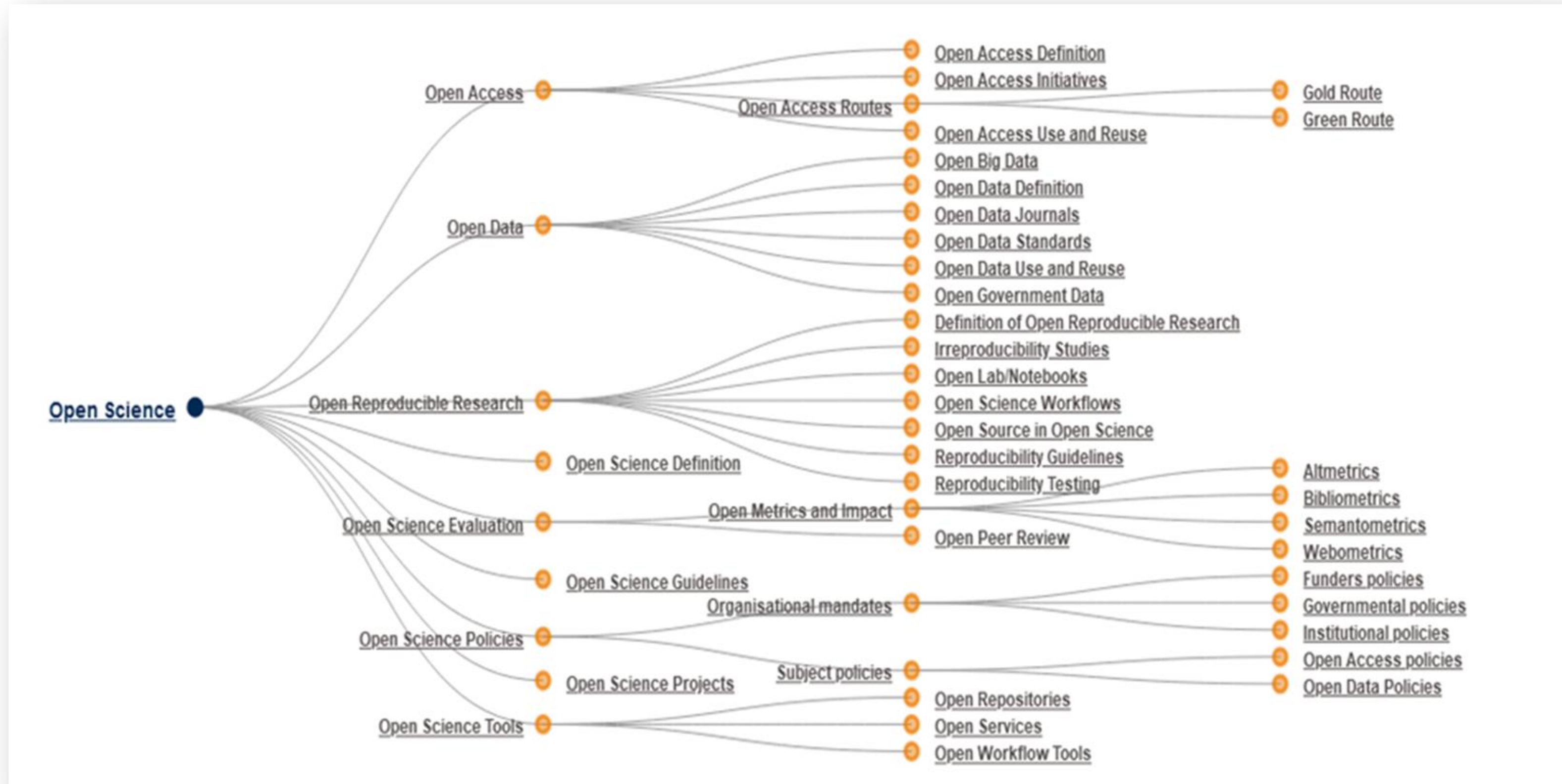
È uno standard di lavoro che si fonda sulla cooperazione e la diffusione della conoscenza.



È finalizzato a migliorare la qualità, l'efficienza e la capacità della ricerca di dare risposte concrete alle sfide poste.



Comprende l'accesso aperto a pubblicazioni e dati, ma non solo!





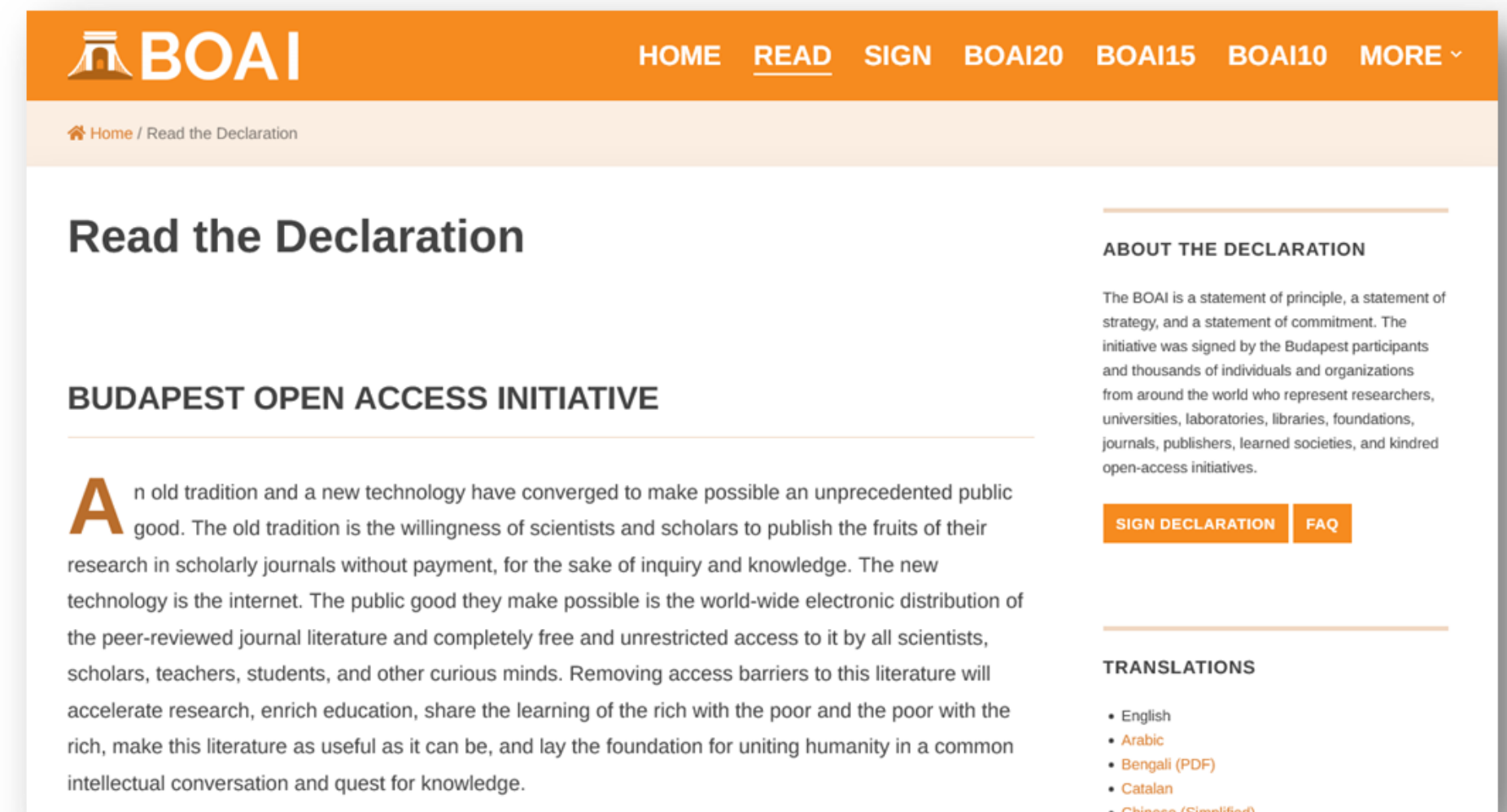
# OPEN SCIENCE: COME CI SIAMO ARRIVATI?



## OPEN SCIENCE: come ci siamo arrivati?

Tutto è iniziato con l'accesso aperto, definito per la prima volta nel 2002 dalla Budapest Open Access Initiative (2002):

La letteratura che dovrebbe essere **liberamente accessibile online** è quella che gli studiosi mettono **a disposizione del mondo senza aspettarsi alcun pagamento** [...]. Per "accesso aperto" a questa letteratura intendiamo la sua libera disponibilità sulla rete Internet pubblica, consentendo a qualsiasi utente di **leggere, scaricare, copiare, distribuire, stampare, ricercare o collegarsi ai testi integrali di questi articoli**, (...), di passarli come dati a un software o di utilizzarli per qualsiasi altro scopo lecito, **senza barriere finanziarie, legali o tecniche** diverse da quelle inseparabili dall'accesso a Internet stesso.



## PERCHÈ OPEN ACCESS?

Le istituzioni si sono rese conto di pagare la ricerca 4 volte:

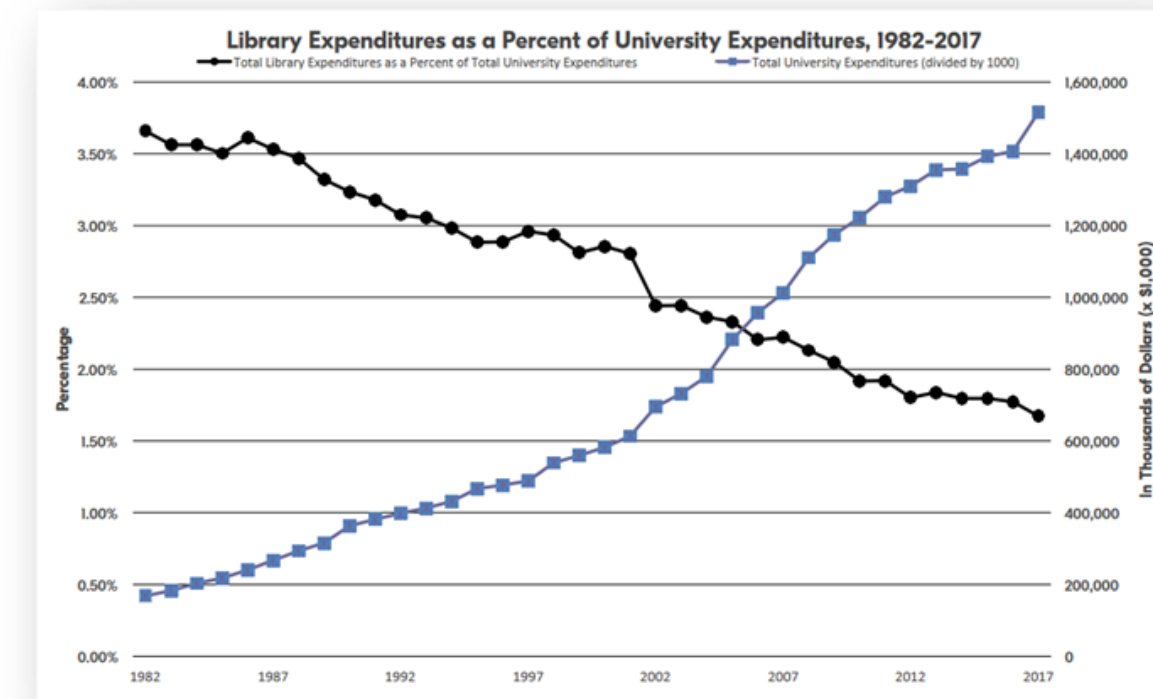
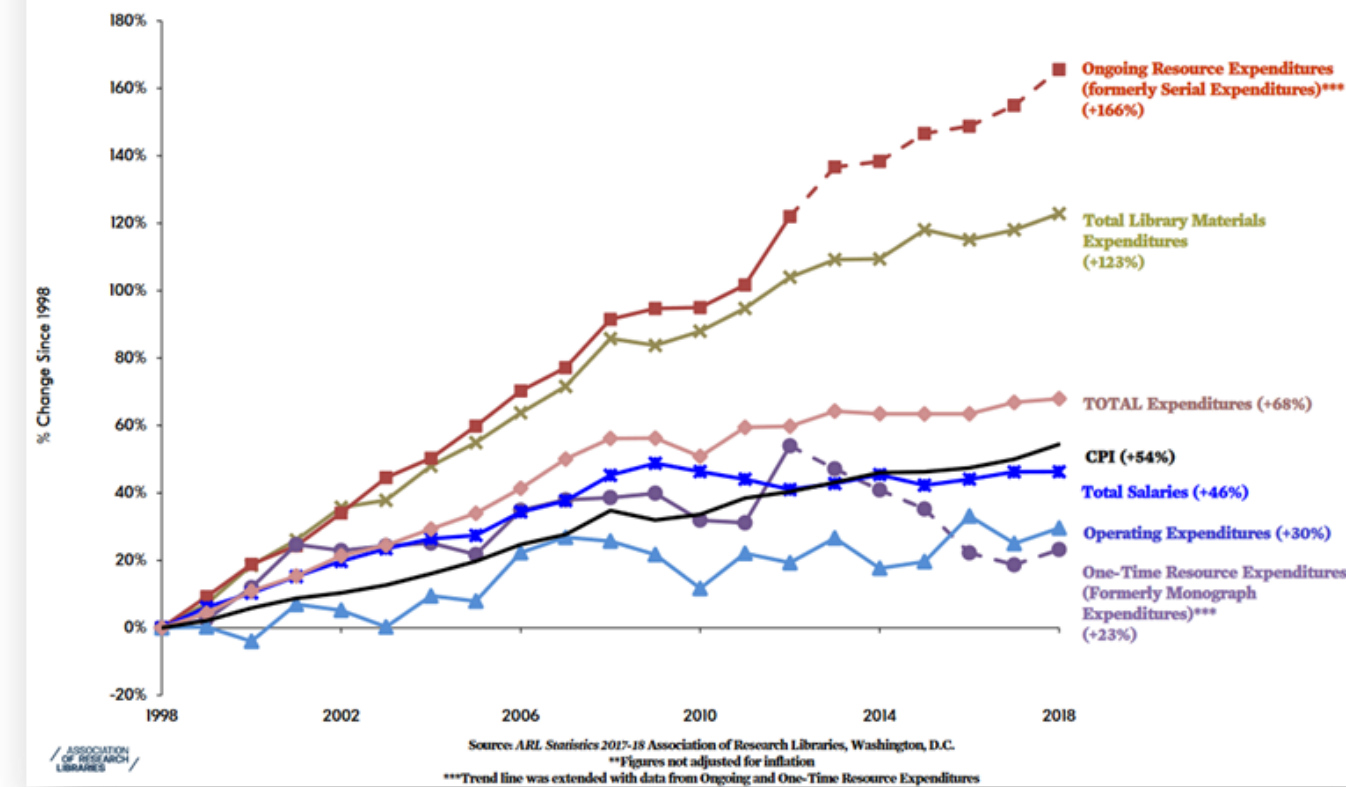
- Stipendi dei ricercatori
- Finanziamento della ricerca
- Abbonamenti a riviste
- Diritti di riutilizzo

Gli autori e i revisori sono pagati dalle università.

Il risultato del lavoro viene rivenduto alle biblioteche dell'Università.

Forte aumento delle spese per i periodici mentre il bilancio della biblioteca continua a essere tagliato

Graph 4  
Expenditure Trends  
in ARL University Libraries, 1998-2018





## DA OPEN ACCESS A OPEN DATA



Solo un anno dopo l'Iniziativa di Budapest per l'accesso aperto, la Dichiarazione di Berlino sull'accesso aperto alla conoscenza nelle scienze e nelle discipline umanistiche (2003) affermava:

Affermare l'accesso aperto come procedura utile richiede idealmente **l'impegno attivo di ogni singolo produttore di conoscenza scientifica e detentore di patrimonio culturale.**

I contributi ad accesso aperto comprendono **i risultati originali della ricerca scientifica, i dati grezzi e i metadati, i materiali di partenza, le rappresentazioni digitali di materiali pittorici e grafici e il materiale multimediale scientifico.**



---

## DA OPEN DATA A OPEN SCIENCE...

By *scholarly communication* we mean the processes of producing, reviewing, organising, disseminating and preserving scholarly knowledge<sup>1</sup>. Scholarly communication does not only concern researchers, but also society at large, especially students, educators, policy makers, public administrators, funders, librarians, journalists, practitioners, publishers, public and private organisations, and interested citizens.

Nel 2016 (con molte tappe intermedie) sono stati pubblicati i Principi di Vienna:

L'attenzione si sposta sulla comunicazione scientifica in generale.

Esprime la convinzione che "la scienza aperta ha il potenziale per cambiare in meglio il funzionamento del sistema di comunicazione accademica".


## ...IL PUNTO IN CUI CI TROVIAMO OGGI

UNESCO Recommendation on Open Science (2021)

<https://unesdoc.unesco.org/ark:/48223/pf0000379949/PDF/379949eng.pdf.multi>



### UNESCO Recommendation on Open Science



EUROPEAN UNION

**HORIZON EUROPE**

#HorizonEU

June 2021

*"Horizon Europe will set a new standard for dissemination of knowledge and new skills across European societies. With clear and immediate open access requirements for beneficiaries, the Open Europe Research publishing platform and a strengthened European Open Science Cloud, we are well underway in making truly open science a reality."*

Mariya Gabriel Commissioner for Innovation, Research, Culture, Education and Youth

**OPEN SCIENCE**

**EARLY KNOWLEDGE AND DATA SHARING, AND OPEN COLLABORATION**

The challenge is for Europe to embrace open science as the modus operandi for all researchers. Open science consists in the sharing of knowledge, data and tools as early as possible in the Research and Innovation (R&I) process, in open collaboration

EC, Directorate-General for R&I, Horizon Europe, open science : early knowledge and data sharing, and open collaboration, 2021,

<https://data.europa.eu/doi/10.2777/18252>



EC, Research, Innovation and Science Policy Experts (2018)

<https://data.europa.eu/doi/10.2777/348700>

---

## PERCHÈ È VANTAGGIOSO PARLARE DI OPEN SCIENCE?



Open Science è un argomento attuale in diversi contesti di finanziamento della ricerca.



La ricerca e i ricercatori si muovono in un contesto internazionale:

- in Europa le strategie legate all'Open Science sono varie e numerose;
- i vostri colleghi in tutta Europa sono formati anche per rispondere a queste richieste.



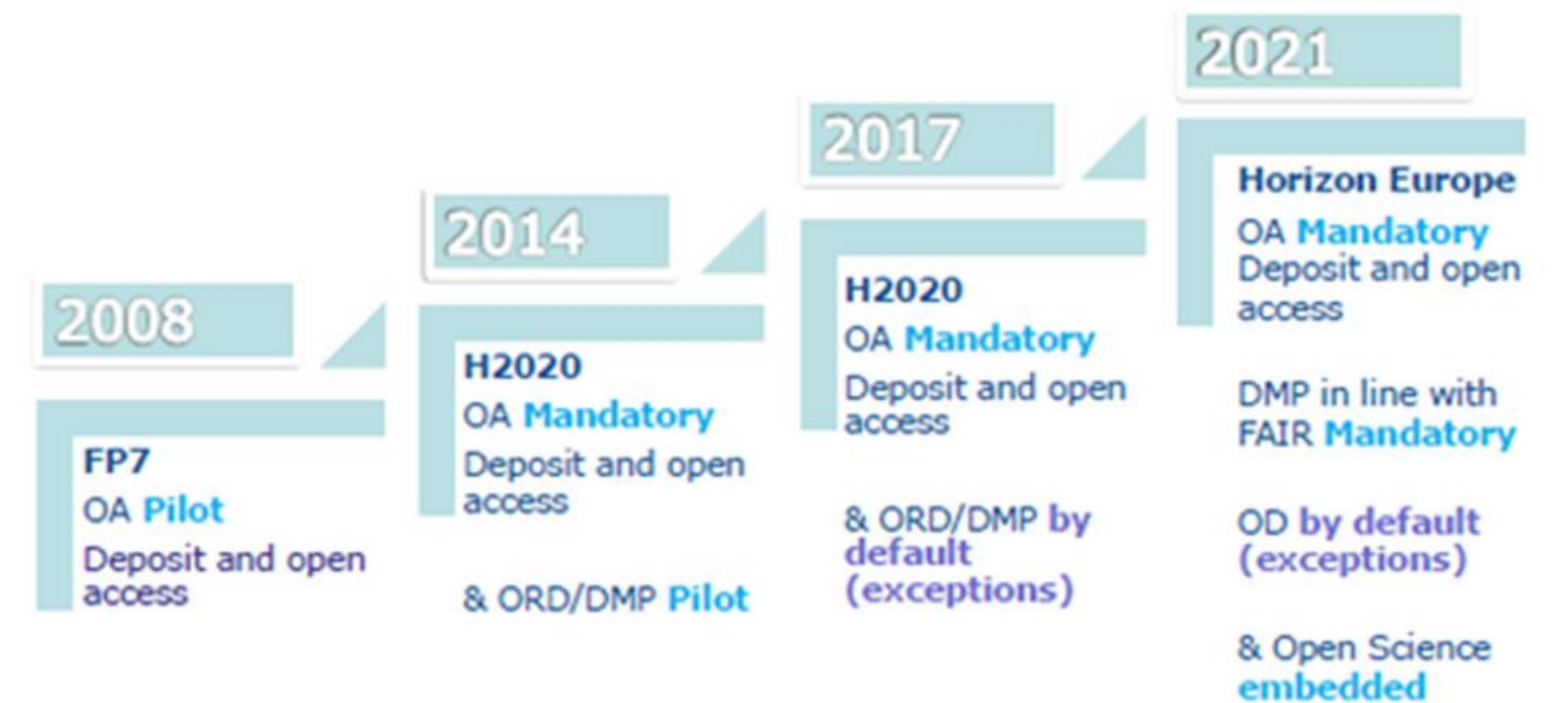
I metodi tradizionali di valutazione della ricerca sono in evoluzione:

- pratiche di Open Science entreranno a far parte dei nuovi criteri.

## LA SCIENZA APERTA NEI FINANZIAMENTI EUROPEI

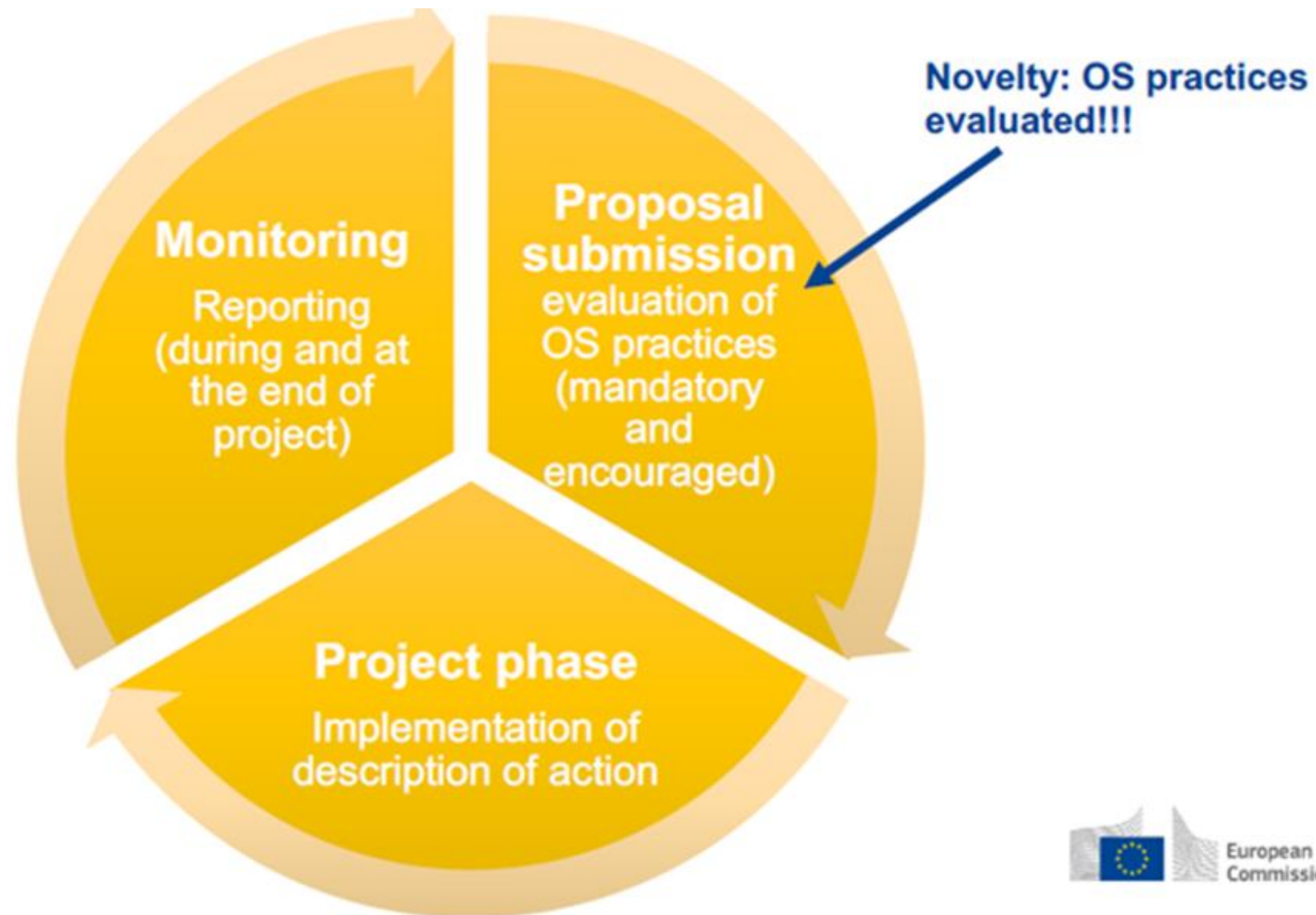
**COSA?** [...] la condivisione di conoscenze, dati e strumenti il più presto possibile nel processo di ricerca e innovazione (R&I), in collaborazione aperta con tutti gli attori della conoscenza, comprese le università, l'industria, le autorità pubbliche, gli utenti finali, i cittadini e la società in generale.

**PERCHÈ?** ha il potenziale per aumentare la qualità, l'efficienza e l'impatto della R&I, portare a una maggiore capacità di risposta alle sfide sociali e aumentare la fiducia della società nel sistema scientifico. Un obiettivo strategico dal 2015,





## LA SCIENZA APERTA NEI FINANZIAMENTI EUROPEI



### Pratiche obbligatorie:

- accesso aperto alle pubblicazioni
- gestione dei dati e degli output della ricerca in linea con i principi FAIR
- DMP, accesso aperto ai dati "il più possibile aperto, il più possibile chiuso".
- accesso digitale o fisico ai risultati della ricerca in caso di emergenza pubblica, accesso aperto immediato a tutti gli output

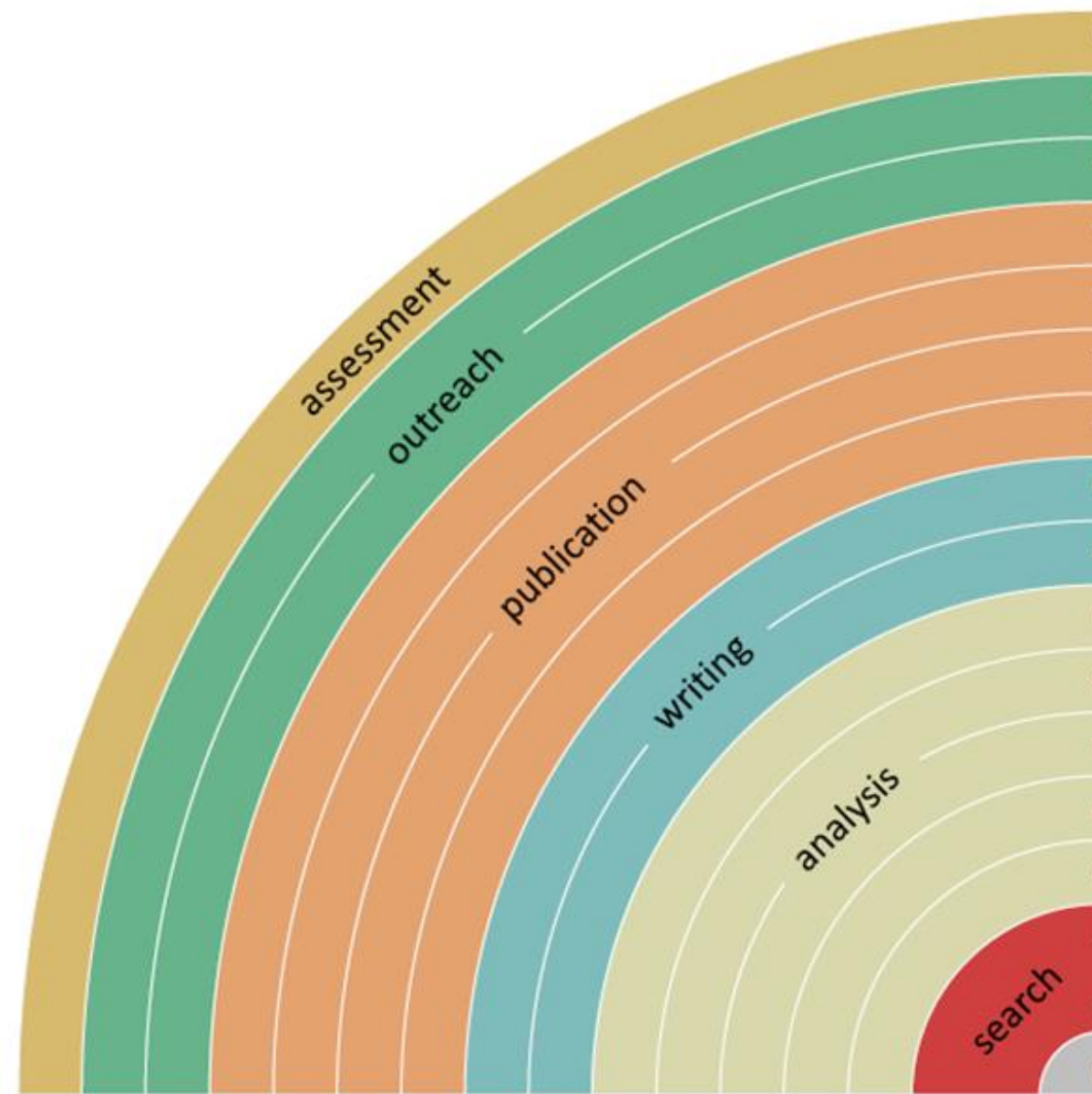
### Pratiche Raccomandate:

- strumenti e pratiche per migliorare la collaborazione, la trasparenza e l'integrità della ricerca
- gestione del flusso di ricerca, collaborazione, riproducibilità
- co-creazione con gli stakeholder e gli utenti finali



## LE PRATICHE DI OPEN SCIENCE SONO VARIE E NUMEROSE

# You can make your workflow more open by ...



- adding alternative evaluation, e.g. with altmetrics
- communicating through social media, e.g. Twitter
- sharing posters & presentations, e.g. at FigShare
- using open licenses, e.g. CC0 or CC-BY
- publishing open access, 'green' or 'gold'
- using open peer review, e.g. at journals or PubPeer
- sharing preprints, e.g. at OSF, arXiv or bioRxiv
- using actionable formats, e.g. with Jupyter or CoCalc
- open XML-drafting, e.g. at Overleaf or Authorea
- sharing protocols & workfl., e.g. at Protocols.io
- sharing notebooks, e.g. at OpenNotebookScience
- sharing code, e.g. at GitHub with GNU/MIT license
- sharing data, e.g. at Dryad, Zenodo or Dataverse
- pre-registering, e.g. at OSF or AsPredicted
- commenting openly, e.g. with Hypothes.is
- using shared reference libraries, e.g. with Zotero
- sharing (grant) proposals, e.g. at RIO



Bianca Kramer & Jeroen Bosman <https://101innovations.wordpress.com>

DOI: [10.5281/zenodo.1147025](https://doi.org/10.5281/zenodo.1147025)



# IL CURRICULUM NARRATIVO: DOVE DARE CONTO DI QUANTO SI LAVORA "OPEN"

*List of up to 5 publications, widely-used datasets, software, goods, services, or any other achievements relevant to the call content.*

| Type of achievement   | Short description  |
|---|--|
| [Publication]<br>[Dataset]<br>[Software]<br>[Good]<br>[Service] | Key elements of the achievement, including a short qualitative assessment of its impact and (where available) its digital object identifier (DOI) or other type of persistent identifier (PID).<br><br>Publications, in particular journal articles, are expected to be open access. Datasets are expected to be FAIR and 'as open as possible, as closed as necessary'. |

## Narrative CV format

The narrative CV is based on the premise that there is no ideal type of researcher. Different research projects require varying talents. The narrative format allows applicants to highlight their own academic profile without being limited by prescribed questions and criteria. The narrative CV consists of two parts:

- The academic profile (narrative)
- Key outputs

The key outputs section focuses on the candidate's past performance:

- A broad definition of 'output' is used. This includes not only scholarly articles but also other outputs, such as datasets, patents and software.
- The number of outputs has been maximized to 10 to stress that NWO values quality over quantity.
- Special attention is paid to Open Science by asking applicants to indicate which of their outputs are openly available.

The narrative CV has been introduced as a pilot in the Veni (part of [NWO Talent Programme](#)) and is gradually being adopted in other NWO funding instruments.

Fase di proposal dei progetti Horizon Europe, considerati per valutazione:

- esperienza in termini di Open Science
- non solo pubblicazioni
- pubblicazioni in Open Access/ dataset FAIR
- no Impact Factor per valutare la qualità delle pubblicazioni

Esempio Olandese: il Dutch Research Council, ente finanziatore nazionale olandese, ha adottato il CV Narrativo per valutare le applicazioni dei ricercatori ai bandi di finanziamento (<https://www.nwo.nl/en/dora>)

# LA RIFORMA DELLA VALUTAZIONE DELLA RICERCA

## Coalition for Advancing Research Assessment

Our vision is that the assessment of research, researchers and research organisations recognises the diverse outputs, practices and activities that maximise the quality and impact of research. This requires basing assessment primarily on qualitative judgement, for which peer review is central, supported by responsible use of quantitative indicators.

**I believe in a research culture that recognises a diversity of contributions to science and society; that celebrates high quality and impactful research; and that values sharing, collaboration, integrity and engagement with society, transmitting knowledge from generation to generation.**

**Mariya Gabriel**

Commissioner for Innovation, Research, Culture, Education and Youth



Italian National Agency for  
the Evaluation of Universities  
and Research Institutes  
(ANVUR)

[website](#) →

## ANVUR firmataria dell'Agreement on reforming research assessment

📅 06/10/2022

ANVUR ha sottoscritto in data odierna l' **"Agreement on reforming research assessment"**, che costituisce il documento programmatico iniziale per una discussione sulla ridefinizione delle strategie e metodologie per la valutazione delle attività di ricerca, da condurre attraverso un percorso comune e condiviso a livello internazionale. In questo senso la partecipazione dell'Agenzia alla Coalition internazionale che si sta definendo è un ulteriore elemento che testimonia la forte volontà dell'Agenzia di partecipare attivamente alle iniziative di respiro internazionale su un tema centrale della sua attività. La decisione, assunta dal Consiglio Direttivo dell'ANVUR nella seduta del 3 ottobre u.s., rappresenta un punto di partenza per le attività che l'Agenzia dovrà sviluppare per la definizione di regole comuni a livello internazionale a partire da questo documento programmatico. La firma giunge al termine di un percorso di partecipazione attiva e di confronto intrapreso attraverso la partecipazione del prof. Menico Rizzi, componente del Consiglio Direttivo con delega alla ricerca sulla valutazione e rappresentante dell'Agenzia in seno alla Coalition internazionale.



# IL PROGRAMMA NAZIONALE PER LA RICERCA

|  |    |
|--|----|
| <b>SOMMARIO</b>  |    |
| <b>PERCHÉ LA SCIENZA APERTA .....</b>  |    |
| <b>IL CONTESTO .....</b>   |    |
| L'apertura come paradigma scientifico .....                                    |    |
| Gli strumenti per l'attuazione della scienza aperta .....                      |    |
| L'impatto della scienza aperta .....   |    |
| <b>STRUTTURA E OBIETTIVI DEL PIANO .....</b>                                   |    |
| <b>ASSI DI INTERVENTO .....</b>  |    |
| <b>1. LE PUBBLICAZIONI SCIENTIFICHE .....</b>                                  |    |
| Il panorama .....  | 5  |
| Il piano di intervento .....   | 6  |
| <b>2. I DATI DELLA RICERCA SCIENTIFICA .....</b>                               | 7  |
| Il panorama .....  | 7  |
| Il piano di intervento .....   | 9  |
| <b>3. LA VALUTAZIONE DELLA RICERCA .....</b>                                   | 10 |
| Il panorama .....  | 10 |
| Il piano di intervento .....   | 10 |
| <b>4. SCIENZA APERTA, COMUNITA' SCIENTIFICA E PARTECIPAZIONE EUROPEA .....</b> | 11 |
| Il panorama .....  | 11 |
| Il piano di intervento .....   | 12 |
| <b>5. APERTURA DEI DATI DELLA RICERCA SU SARS-COV-2 E COVID-19 .....</b>       | 12 |
| Il panorama .....  | 12 |
| Il piano di intervento .....   | 13 |
| <b>GLOSSARIO .....</b>   | 14 |
| <b>DEFINIZIONI .....</b>   | 16 |
| <b>DOCUMENTI DI RIFERIMENTO .....</b>  | 16 |
| <b>RIFERIMENTI NORMATIVI .....</b>   | 19 |



**PNR** 2021-2027  
**Programma nazionale per la ricerca**



# OPEN SCIENCE NEL PIANO STRATEGICO DI UNIBO



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

## PIANO STRATEGICO 2022-2027

[Home](#) [Mission e vision](#) **[Principi](#)** [Obiettivi](#) [It](#) [En](#)

[Home](#) > [Principi](#) > [Incrementare il senso della nostra responsabilità sociale in tutte le attività](#)



OBIETTIVO 31

### Favorire lo sviluppo dell'open science

**P3** PRINCIPIO 3

**Q** RICERCA



# QUALCHE PUNTO CHIAVE

- Open Science è uno standard di lavoro che si fonda sulla cooperazione e la diffusione della conoscenza.
- Contiene in sé principi di trasparenza, integrità della ricerca, collaborazione piena tra attori diversi, libertà di accesso a dati e pubblicazioni, nuove metriche di valutazione.
- E' finalizzato a migliorare la qualità, l'efficienza e la capacità della ricerca di dare risposte concrete alle sfide poste









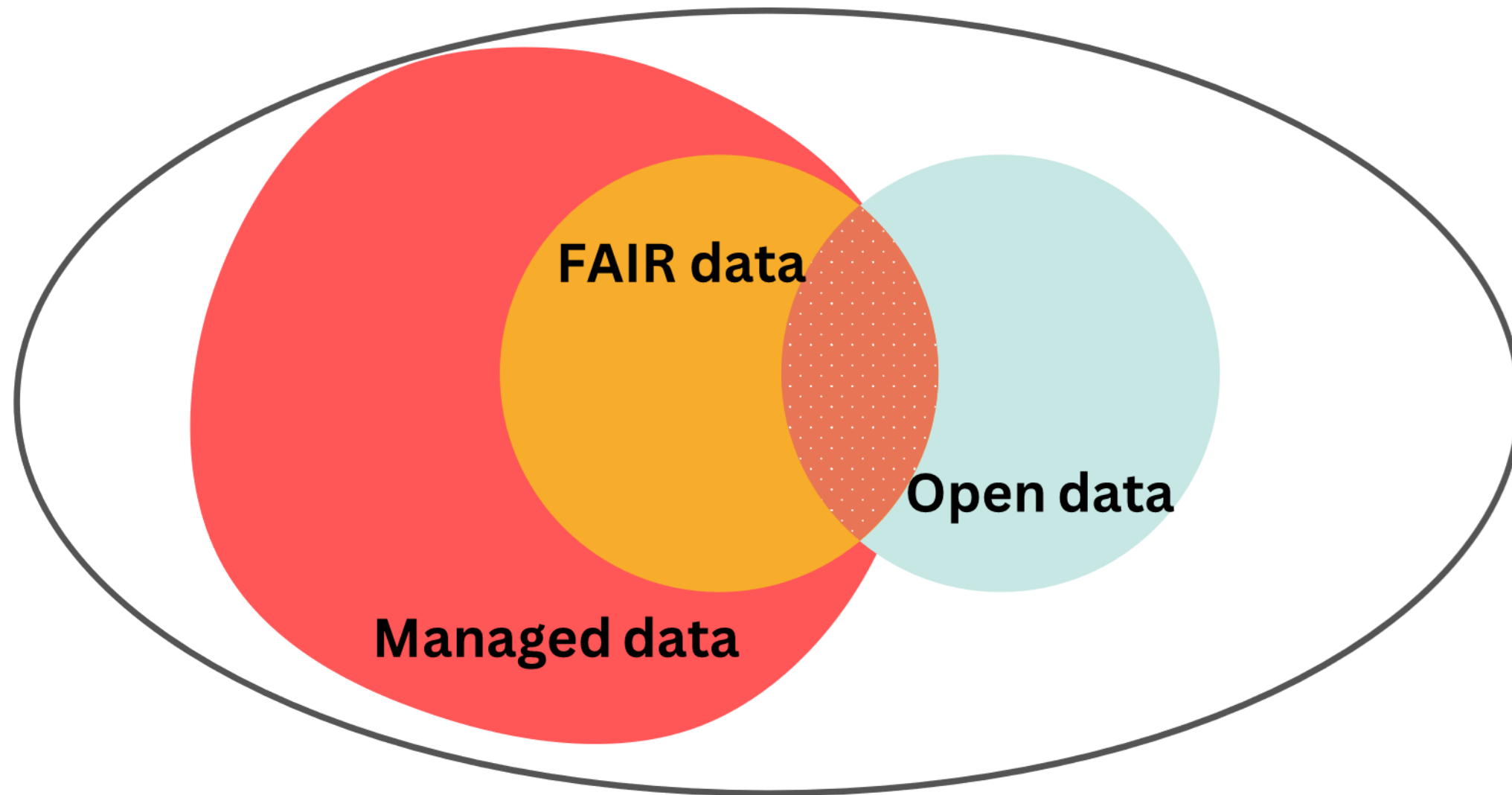


# I PRINCIPI FAIR





## FAIR: COS'È?



- I principi FAIR descrivono come i risultati della ricerca dovrebbero essere organizzati per essere più facilmente accessibili, compresi, scambiati e riutilizzati.
- Non tutti i dati gestiti correttamente sono FAIR.
- Non tutti i dati FAIR sono dati Open.

## FAIR: 15 PRINCIPI GUIDA

# SCIENTIFIC DATA

Amended: Addendum

**OPEN**

SUBJECT CATEGORIES

- » Research data
- » Publication characteristics

Received: 10 December 2015  
Accepted: 12 February 2016  
Published: 15 March 2016

## Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*<sup>#</sup>

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data by a diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers. We have come together to design and jointly endorse a concise and measurable set of principles to guide the management and stewardship of scientific data. These principles, known as the FAIR Data Principles, are intended to act as a guideline for those who manage and steward data, to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on data sharing, the FAIR Principles put specific emphasis on enhancing the ability of machines to find and use the data, in addition to supporting its reuse by individuals. This Comment represents the formal publication of the FAIR Principles, and includes the rationale behind them, and some initial implementations in the community.

### Box 2 | The FAIR Guiding Principles

**To be Findable:**

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

**To be Accessible:**

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
  - A1.1 the protocol is open, free, and universally implementable
  - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

**To be Interoperable:**

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

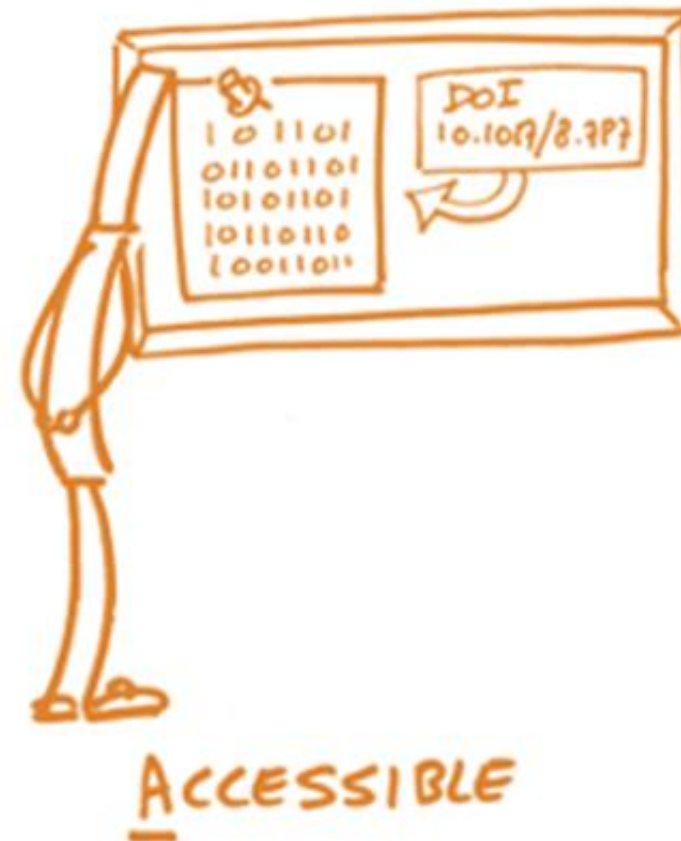
**To be Reusable:**

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
  - R1.1. (meta)data are released with a clear and accessible data usage license
  - R1.2. (meta)data are associated with detailed provenance
  - R1.3. (meta)data meet domain-relevant community standards

## FAIR DATA PRINCIPLES



Il primo passo per (ri)utilizzare i dati è trovarli.



Una volta trovati, l'utente deve sapere come accedere (autenticazione, autorizzazione...)



I dati devono essere integrati con altri dati e devono interoperare con applicazioni o flussi di lavoro.



L'obiettivo finale di FAIR è ottimizzare il riutilizzo dei dati.



## FAIR FOR HUMANS AND MACHINES



### 1) Both humans and machines are intended as digesters of data.

This will lead to the creation of an ecosystem that is fast to respond to change and automatically adapts to new findings or changes: the *Internet of FAIR Data and Services*. This is the reason for focusing on standards for data, identification mechanisms, data availability, etc.

F indable

A ccessible

I nteroperable

R eusable

Per gli esseri  
umani e per le  
macchine

F ully

A (rtificial)

I (ntelligence)

R eady



---

## PUNTI DI ATTENZIONE PER APPLICARE I PRINCIPI FAIR

- I principi FAIR si applicano sia ai dati che ai metadati.
- I principi non riguardano necessariamente i dati aperti («as open as possible, as closed as necessary»).
- I principi FAIR non sono regole o standard, ma un adattamento delle routine di lavoro.

## COSA INTENDIAMO QUANDO PARLIAMO DI METADATI?

Informazioni strutturate sui dati

- Li descrivono, spiegano, localizzano
- Ne facilitano l'uso o la gestione
- Possono essere aggiunti manualmente o automaticamente
- Esistono standard specifici per disciplina



# COSA INTENDIAMO QUANDO PARLIAMO DI METADATI?

## METADATI



[Questa foto](#) di Autore sconosciuto è concesso in licenza da [CC BY-SA](#)

### INFORMAZIONI SUI DATI:

- Autore
- Titolo
- Progetto
- Sigla
- Ecc

## DATI

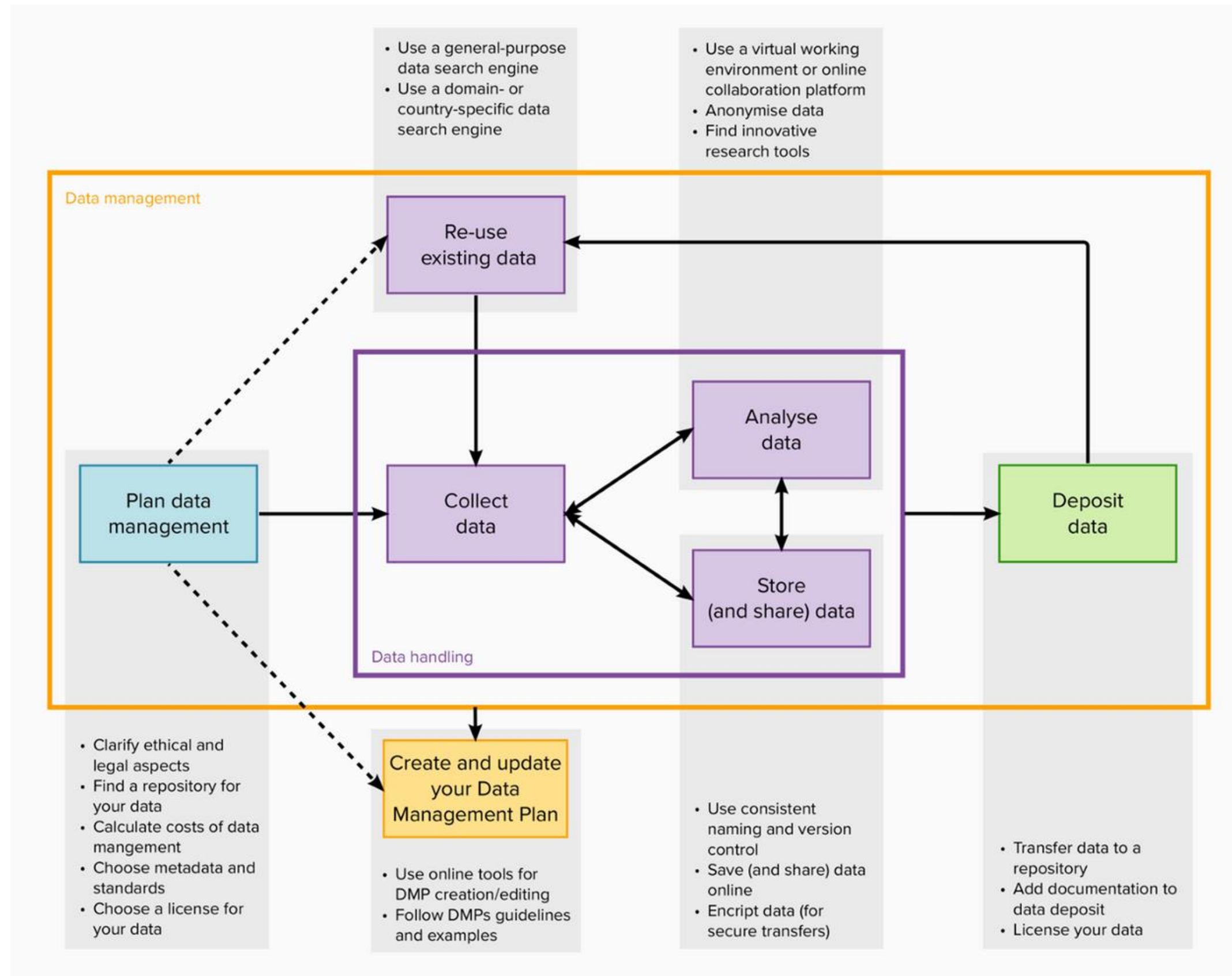


[Questa foto](#) di Autore sconosciuto è concesso in licenza da [CC BY-NC](#)

|                                  |  |
|----------------------------------|--|
| Showing Metadata for:            | Country ▾  |
| Code                             | ITA  |
| Long Name                        | Italian Republic   |
| Income Group                     | High income  |
| Region                           | Europe & Central Asia  |
| Other groups                     | Euro area  |
| Currency Unit                    | Euro   |
| Latest population census         | 2012. Population figures compiled from administrative registers in combination with other sources of data, such as sample surveys.   |
| Special Notes                    | A simple multiplier is used to convert the national currencies of EMU members to euros. The following irrevocable euro conversion rate was adopted by the EU Council on January 1, 1999: 1 euro = 1936.27 Italian lira. Please note that historical data before 1999 are not actual euros and are not comparable or suitable for aggregation across countries. |
| National accounts base year      | Original chained constant price data are rescaled.   |
| National accounts reference year | 2015   |
| System of National Accounts      | Country uses the 2008 System of National Accounts methodology  |
| SNA price valuation              | Value added at basic prices (VAB)  |
| PPP survey years                 | Rolling surveys (annual estimation)  |

[https://databank.worldbank.org/source/world-development-indicators#selectedDimension\\_WDI\\_Ctry](https://databank.worldbank.org/source/world-development-indicators#selectedDimension_WDI_Ctry)  
CC BY

## I PRINCIPI FAIR SONO TRASVERSALI: DA APPLICARE DURANTE L'INTERO CICLO DI VITA DEL DATO







F  
indable

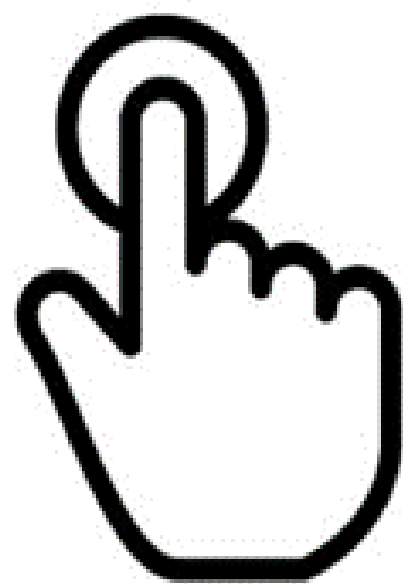


- Ai dataset devono essere assegnati identificatori persistenti (ad esempio, DOI, Handle, ecc).
- I dataset devono essere accompagnati da metadati e parole chiave significative (contenenti il PID).



# A

ccessible





- I metadati devono essere apertamente accessibili attraverso l'uso di protocolli standard.
- I dati possono essere apertamente accessibili (default) o accessibili attraverso un sistema di autenticazione e autorizzazione, se la natura dei dati ne impedisce l'apertura.
- I dati e i metadati devono essere conservati in archivi o depositi che li rendano persistenti nel tempo e rintracciabili in rete.

---

## "AS OPEN AS POSSIBLE AS CLOSED AS NECESSARY"

- Lavorate con dati personali? Seguire il GDPR e proteggere la privacy
- Lavorate con dati sensibili? Seguire il GDPR, potrebbero essere coinvolti i comitati etici.
- Volete riutilizzare i dati in base ai diritti di proprietà intellettuale o al copyright?
- Volete brevettare? Prevedere un embargo sui dati
- Volete pubblicare (ad esempio, una monografia) e temete che le idee possano essere rubate? Si può imporre un embargo e/o pubblicare i dati sottostanti insieme alla pubblicazione.

- 
- Consenso informato
  - Anonimizzazione
  - Accesso controllato
  - Comitati di accesso



Controllare le licenze! Il riutilizzo può essere limitato, poiché l'accessibilità dei nuovi risultati è limitata.

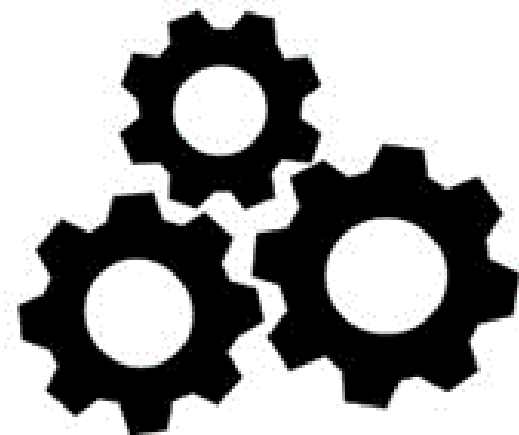
- 
- Deposita i dati con embargo
  - Accesso controllato





# I

## nteroperable

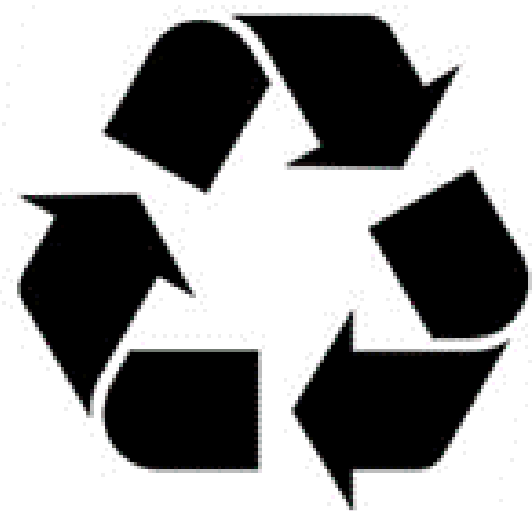


- I dati devono essere combinabili e utilizzabili con altri dati o strumenti.
- Il formato dei dati deve essere aperto e interoperabile da vari strumenti, compresi altri database.
- I metadati devono utilizzare un linguaggio standardizzato e condiviso a livello internazionale dai diversi servizi di indicizzazione.
- Includere riferimenti qualificati ad altri (meta)dati.



# R

Reusable



- I dati devono essere descritti e documentati nel miglior modo possibile, per garantirne la qualità e per poter essere replicati e combinati in contesti diversi.
- L'elaborazione dei dati deve essere conforme agli standard riconosciuti dalle comunità scientifiche di riferimento.
- Il riutilizzo dei (meta)dati deve essere dichiarato sotto una o più licenze aperte chiare e accessibili (ad esempio, Creative Commons, GPL per il software e altre).



# COME VALUTARE SE UN DATASET È FAIR?

---

TESTIAMO UN DATASET INSIEME:



F-UJI is a web service to programmatically assess FAIRness of research data objects at the dataset level based on the FAIRsFAIR Data Object Assessment Metrics [↗](#)

[Click here to assess a dataset](#)

<https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>





# QUALCHE PUNTO CHIAVE

- I dati della ricerca devono essere gestiti in modo trasparente: seguire i principi FAIR e renderli reperibili, accessibili, interoperabili e riutilizzabili.
- I dati FAIR non sono necessariamente dati gestiti correttamente né tantomeno Open Data ☐ "As open as possible, as closed as necessary".
- Un repository per i dati non fa tutto il lavoro per rendere i vostri dati il più FAIR possibile. Tuttavia, offre un'ottima struttura per ottenere le basi giuste.



## CONSIGLI DI APPROFONDIMENTO

### Open Science:

Open Science: A Practical Guide for Early-Career Researchers (1.0) <https://doi.org/10.5281/zenodo.7716153>

Open Science Toolkit (UNESCO) <https://www.unesco.org/en/open-science/toolkit>

Portale Italiano Open-Science <https://open-science.it>

Webinar Open Science Cafè <https://www.icdi.it/it/attivita/competence-centre/open-science-cafe>

### Gestione FAIR dei Dati di Ricerca:

Open AIRE guidelines <https://guidelines.openaire.eu/en/latest/>

The Turing Way handbook <https://the-turing-way.netlify.app/index.html>

The hands-on guide to research data management for KU Leuven researchers, students, and research support staff in the humanities and social sciences. <https://doi.org/10.5281/zenodo.4596323>

LEGO® Metadata for Reproducibility game pack <http://eprints.gla.ac.uk/196477/>

Reprojuice, the reproducibility game <https://seriousgaming.tudelft.nl/game/reprojuice>







GRAZIE

