

# Le basi del Research Data Management

**Perché è fondamentale prendersi cura dei dati**  
**Cisup - 20 gennaio 2022**

Gina Pavone

Isti-Cnr

Emma Lazzeri

Garr

# Gina Pavone

- Research fellow at the Institute of Information Science and Technologies of the Italian National Research Council in Pisa, Italy.
- Research focus: Open Science and Open Access; Research Data Management
- OpenAIRE National Open Access Desk (NOAD) for Italy
- Coordinator of the editorial board of [open-science.it](https://open-science.it) website
- My background: data journalism





# Emma Lazzeri

- Open Science expert at Consortium GARR
- Associate Researcher at Institute of Information Science and Technologies of the Italian National Research Council in Pisa, Italy.
- Research focus: Open Science Strategies and Policies; Research Data Management
- ICDI Competence Centre Coordinator
- My background: PhD in Innovative Technologies





# Di cosa parleremo oggi

Seminario introduttivo sull'importanza del RDM



01

## Perché

Comunicazione scientifica e Open Science. Il problema della valutazione

02

## New normal

Horizon Europe, la scienza aperta e la nuova normalità

03

## RDM

Che cos'è e perché è importante

04

## FAIR

Cosa sono i principi FAIR

05

## CISUP

Elaborare una strategia 'personalizzata' per gestire i dati



# A cosa pensi quando senti parlare di Research Data Management?

[15:09] Angelica Salvadori86912033

828 5644 4869

progetti europei

trattamento dei dati risultanti da analisi

Un sistema per gestire i dati raccolti all'interno di un'attività di ricerca

Progetti di ricerca europei  
Progetti in grandi collaborazioni

gestione dei dati riservati e non

Definizione di un protocollo per la gestione dei dati, dalla loro acquisizione fino alla loro archiviazione e manutenzione

Gestione ragionata dei dati raccolti



# A cosa pensi quando senti parlare di Research Data Management?

non conosco: gestione dei dati di ricerca?

al deliverable a M6

Sono nuova all'argomento, finora nel gruppo lo abbiamo gestito "in casa"

non ho idee chiare

Gestione di dati usati per i progetti, simulazioni

acquisizione, immagazzinamento e gestione dei dati in conformità a specifiche che ne consentano la condivisione e la reperibilità

Attività di gestione volta a non perdere i dati e a renderli riutilizzabili

Generalmente direi all'acquisizione, gestione, utilizzo ed archiviazione/manutenzione dei dati di ricerca derivanti da tutte le possibili fonti.

Archiviazione e conservazione dei dataset raccolti nel corso di un progetto



# A cosa pensi quando senti parlare di Research Data Management?

organizzazione e sicurezza dei dati

Non ne sono informata. Deduco procedure corretta dei dati di ricerca

organizzazione e interpretazione e utilizzo dei dati

come organizzare, analizzare e conservare i dati della ricerca

Come gestire la mole di dati ottenuti in una ricerca. Come organizzarli e gestirli per un lavoro ottimale, ed un eventuale paper.

gestione dei dati di ricerca, non so come?

Razionale organization of data to boost scientific insight

Modalità di condivisione armonizzazione di conservazione database , metodi per l'interpretazione di dati (multivariati) e per trasformare risultati in risposte

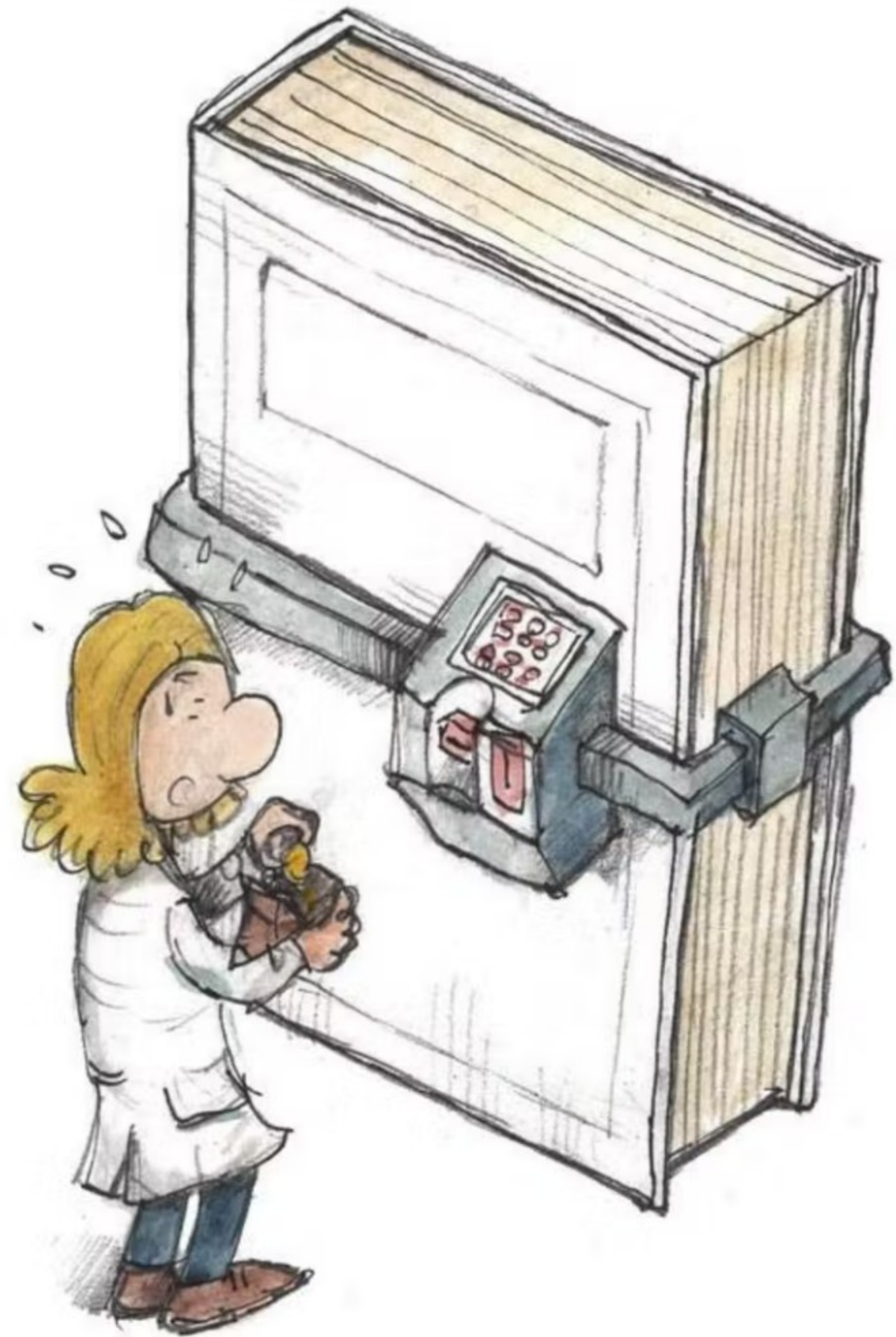


Perché l'Open Science



# La conoscenza, dietro un paywall

Uno dei problemi del sistema di  
comunicazione scientifica  
tradizionale





[nature](#) > [nature reviews physics](#) > [world view](#) > [article](#)


[World View](#) | [Published: 29 September 2021](#)

# How positioning students in the professional physics community leads to a more inclusive community

[Martha-Elizabeth Baylor](#) 

[Nature Reviews Physics](#) (2021) | [Cite this article](#)

**6** [Accesses](#) | **3** [Altmetric](#) | [Metrics](#)

 [Access through your institution](#)

[Buy or subscribe](#)

## Sections

[Acknowledgements](#)

[Author information](#)

Se vuoi leggere questo articolo bisogna pagare. Direttamente o indirettamente



[nature](#) > [nature reviews physics](#) > [perspectives](#) > [article](#)

Perspective | [Published: 28 September 2021](#)

# Visualizing big science projects

[Katy Börner](#) , [Filipi Nascimento Silva](#) & [Staša Milojević](#)

[Nature Reviews Physics](#) (2021) | [Cite this article](#)

**39** [Accesses](#) | **8** [Altmetric](#) | [Metrics](#)

## Abstract

The number, size and complexity of ‘big science’ projects are growing – as are the

 [Access through your institution](#)

[Buy or subscribe](#)

**Sections**

[Figures](#)

[References](#)

[Abstract](#)

[Code availability](#)

[References](#)

Per articoli con approcci multidisciplinari potremmo non avere l'accesso istituzionale



**La pandemia e  
l'urgenza di  
aprire i  
risultati della  
ricerca**





Perché l'accesso aperto non è il default?





# La vostra istituzione paga per:

- Il lavoro dei ricercatori come **autori**
- Il lavoro dei ricercatori come **revisori**
- Il lavoro dei ricercatori come **editor**
- **L'accesso** ai risultati della ricerca  
(non il possesso degli articoli!)





# 10 miliardi di \$

la stima della spesa globale  
**annuale** per l'accesso alle  
riviste scientifiche su scala  
mondiale.





**I ricercatori cedono gratuitamente i propri diritti d'autore commerciali alle riviste**



# An industry like no other

In 2010, Elsevier's scientific publishing arm reported profits of £724m on just over £2bn in revenue. It was a **36% margin** – higher than Apple, Google, or Amazon posted that year.





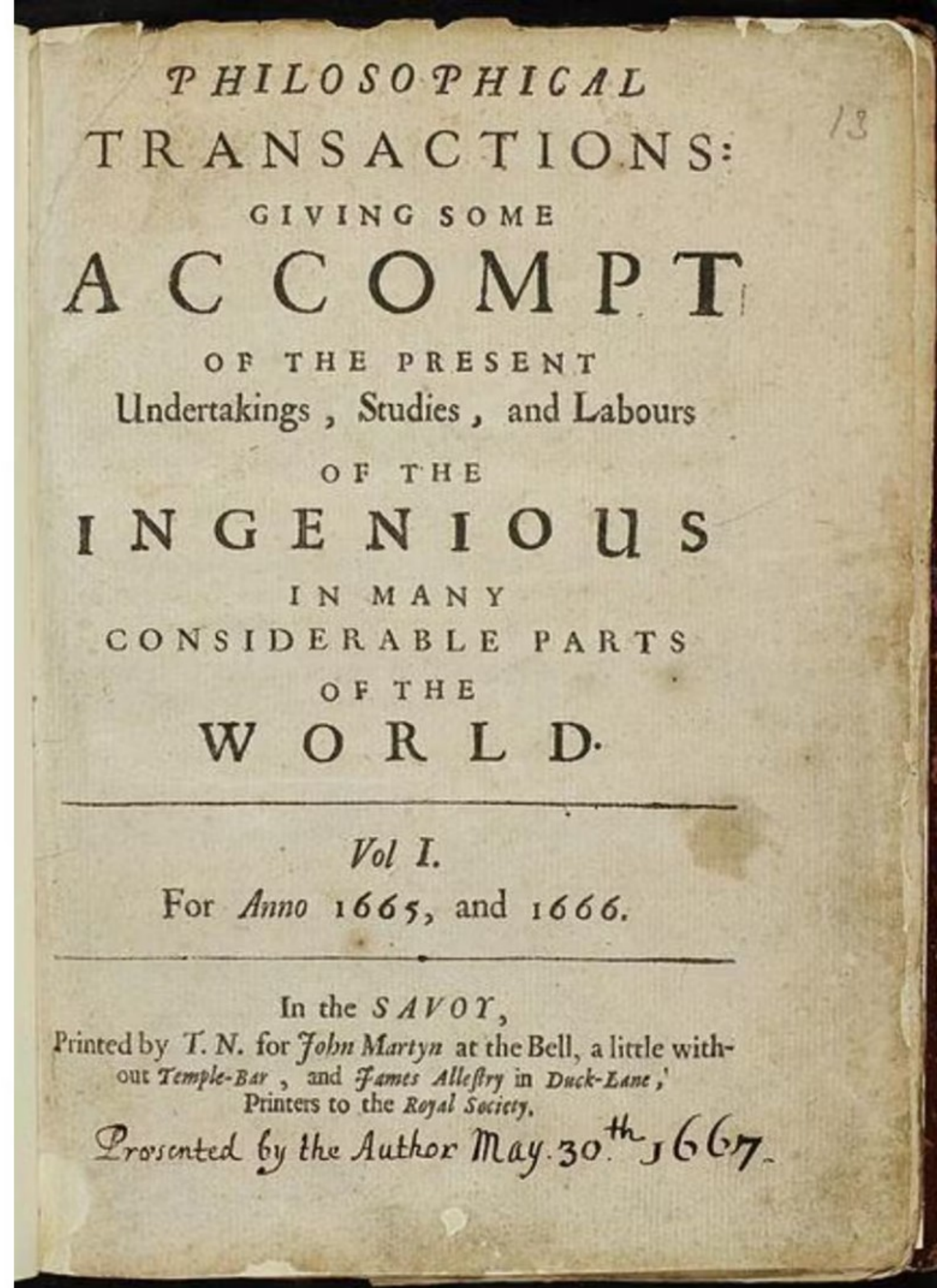
**Come siamo arrivati fino  
a questo punto?**

---



# La nascita delle riviste scientifiche

- 1665, nasce la **prima rivista** interamente dedicata alla scienza
- Obiettivi della prima rivista: migliorare la **comunicazione** tra scienziati, che fino a questo momento avveniva attraverso **lettere** inviate attraverso società private e reti di corrispondenza





# Il copyright è un'invenzione

## Prima della stampa

- I poeti ripetono il canto delle muse
- "Scientia donum dei est, unde vendi non potest"

## Il diritto romano e le res qui tangi possunt

- La scienza è un bene comune

## Il regime dei privilegi

- Autorizzazione esclusiva alla stampa concessa dal potere politico
- L'oggetto non è la proprietà intellettuale ma l'azione della stampa e del commercio dei libri
- Forma di controllo politico sulla produzione e disseminazione della conoscenza (censura)



# La nascita del copyright

- 1710 il parlamento britannico promulga la Statuto di Anna, il primo statuto sul copyright.
- Garantiva protezione legale per 14 anni durante i quali l'autore e il tipografo a cui era stata concesso licenza di stampa potevano pubblicare il lavoro dell'autore.

Anno Octavo  
Annæ Reginae.

An Act for the Encouragement of Learning, by Vesting the Copies of Printed Books in the Authors or Purchasers of such Copies, during the Times therein mentioned.



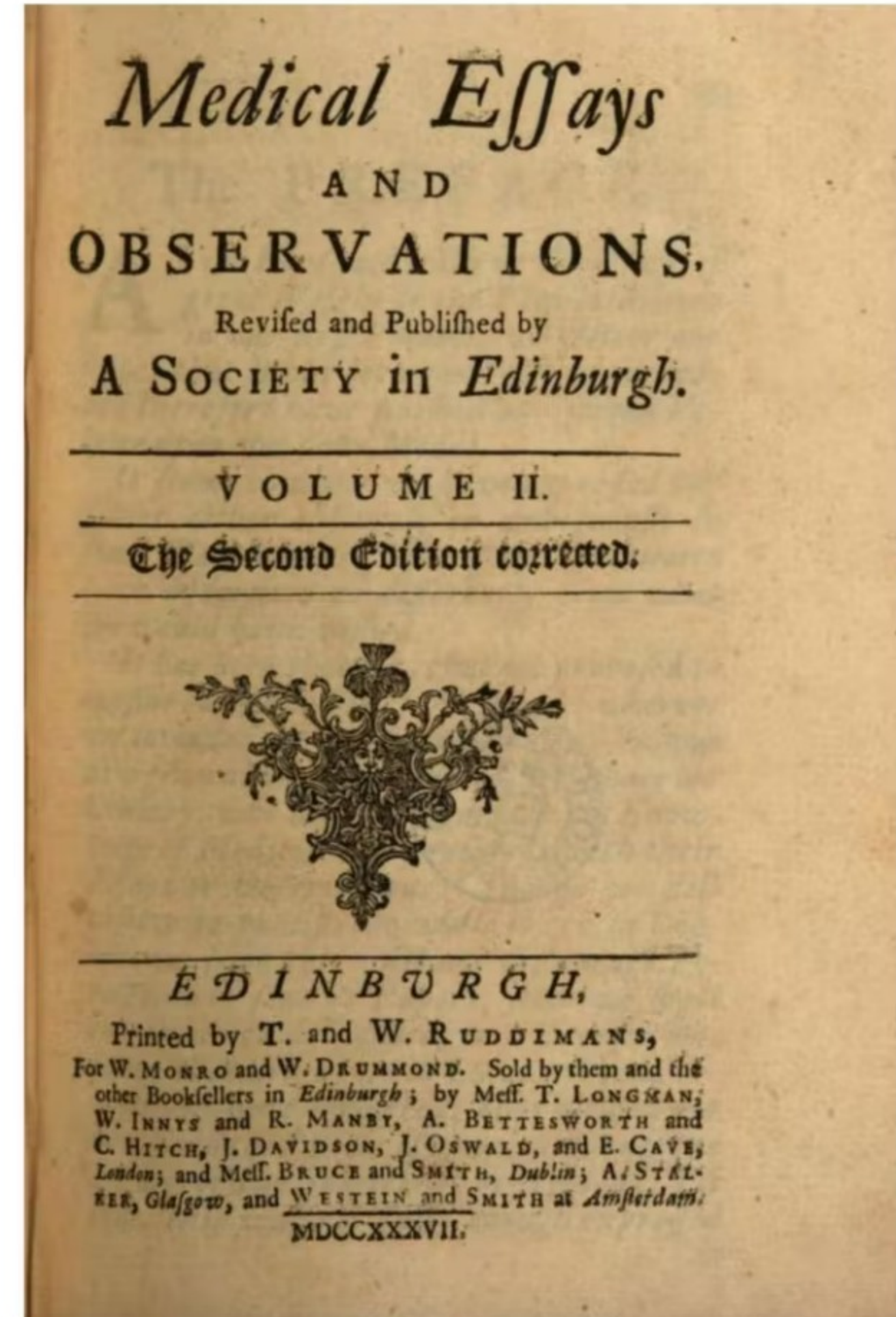
Whereas Printers, Booksellers, and other Persons have of late frequently taken the Liberty of Printing, Reprinting, and Publishing, or causing to be Printed, Reprinted, and Published Books, and other Writings, without the Consent of the Authors or Proprietors of such Books and Writings, to their very great Detriment, and too often to the Ruin of them and their Families: For Preventing therefore such Practices for the future, and for the

Encouragement of Learned Men to Compose and Write useful Books: May it please Your Majesty, that it may be Enacted, and be it Enacted by the Queens most Excellent Majesty, by and with the Advice and Consent of the Lords Spiritual and Temporal, and Commons in this present Parliament Assembled, and by the Authority of the same, That from and after the Tenth Day of April, One thousand seven hundred and ten, the Author of any Book or Books already Printed, who hath not Transferred to any other the Copy or Copies of such Book or Books, Share or Shares thereof, or the Bookseller or Booksellers, Printer or Printers, or other Person or Persons, who hath or have Purchased or Acquired the Copy or Copies of any Book or Books, in order to Print or Reprint the same, shall have the sole Right and Liberty of Printing such Book and Books by the Term of One and twenty Years, to Commence from the said Tenth Day of April, and no longer; and that the Author of any Book or Books already Composed and not Printed and Published, or that shall hereafter be Composed, and his Assigns, or Assigns, shall have the sole Liberty of Printing and Reprinting such Book and Books for the Term of Four-



# La nascita della peer-review

- Un precursore già presente nella prima rivista scientifica (autorizzazione della Royal Society: controllo sui privilegi)
- Nel 1731 *Medical Essays and Observations* pubblicato dalla Royal Society of Edinburgh
- *“Memoirs sent by correspondence are distributed according to the subject matter to those members who are most versed in these matters. The report of their identity is not known to the author.”*





# Comunicazione scientifica

“the system through which research and other scholarly writings are **created, evaluated** for quality, **disseminated** to the scholarly community, and **preserved** for future use”

(ACRL, Principles and Strategies for the Reform of Scholarly Communication 1, 2003).

**Formalised publishing practices are just a subset of a larger pool of various communication practices** (emails, social media, blogs, press, etc.), both between scholars and between scholars and the public.

SC “includes both the dissemination and access to scholarship and research **in a variety of formats and states of completion**, such as published books or journal articles, research results and data sets, and drafts of papers”

(S. Husain S, M. Nazim, Analysis of Open Access Scholarly Journals in Media & Communication. DESIDOC Journal of Library and Information Technology. 2013; 33(5), p. 405–11)



# Il kit di sopravvivenza dei ricercatori

- Pubblicare il più possibile
- Avere un alto numero di citazioni
- “importanti” sedi (alto IF)

Tutto questo non implica necessariamente che la ricerca sia di qualità!

19th century  
scientist

I must find the explanation for this phenomenon in order to truly understand Nature...



21st centurt  
~~scientist~~  
academic

I must get the result that fits my narrative so I can get my paper into Nature..







**Integrità della ricerca: abbiamo un problema.**



“It's natural to think of scientists as truth seekers, people driven by an intense curiosity to understand the natural world.

Yet this picture of scientists and scientific inquiry sits uncomfortably with the reality and prevalence of scientific fraud. If one wants to get at the truth about nature, why lie?”

(2021). LIAM KOFI BRIGHT, Why Do Scientists Lie? Royal Institute of Philosophy Supplement. 89. 117-129. 10.1017/S1358246121000102.



## Alcuni casi famosi

Brian Wansink - Ricercatore in scienze della nutrizione alla Cornell University - 42 pubblicazioni ritirate in 25 riviste scientifiche



Yoshitaka Fujii - Anestesiologo alla Toho University - ha inventato i dati in 171 articoli scientifici





Non sono i soli...

[www.retractionwatch.com](http://www.retractionwatch.com)

## Retraction Watch

Tracking retractions as a window into the scientific process

### PAGES

[How you can support Retraction Watch](#)

[Meet the Retraction Watch staff](#)

[About Adam Marcus](#)

[About Ivan Oransky](#)

[Papers that cite Retraction Watch](#)

[Privacy policy](#)

[Retracted coronavirus \(COVID-19\) papers](#)

[Retraction Watch Database User Guide](#)

[Retraction Watch Database User Guide Appendix A: Fields](#)

[Retraction Watch Database User Guide Appendix B: Reasons](#)

[Retraction Watch Database User Guide Appendix C: Article Types](#)

[Retraction Watch Database User Guide Appendix D: Changes](#)

[The Center For Scientific Integrity](#)

[Board of Directors](#)

[The Retraction Watch FAQ, including comments policy](#)

[The Retraction Watch Transparency Index](#)

[The Retraction Watch Leaderboard](#)

[Top 10 most highly cited retracted papers](#)

## The Retraction Watch Leaderboard

Who has the most retractions? Here's our unofficial list (see notes on methodology), which we'll update as more information comes to light:

1. [Yoshitaka Fujii](#) (total retractions: 183) See also: [Final report of investigating committee](#), [our reporting](#), [additional coverage](#)
2. [Joachim Boldt](#) (159) See also: [Editors-in-chief statement](#), [our coverage](#)
3. [Yoshihiro Sato](#) (106) See also: [our coverage](#)
4. [Ali Nazari](#) (85) See also: [our coverage](#)
5. [Jun Iwamoto](#) (82) See also: [our coverage](#)
6. [Diederik Stapel](#) (58) See also: [our coverage](#)
7. [Yuhji Saitoh](#) (53) See also: [our coverage](#)
8. [Adrian Maxim](#) (48) See also: [our coverage](#)
9. [Chen-Yuan \(Peter\) Chen](#) (43) See also: [SAGE](#), [our coverage](#)
10. [Fazlul Sarkar](#) (41) See also: [our coverage](#)
11. [Shahaboddin Shamshirband](#) (41) See also: [our coverage](#)
12. [Hua Zhong](#) (41) See also: [journal notice](#)
13. [Shigeaki Kato](#) (40) See also: [our coverage](#)
14. [James Hunton](#) (37) See also: [our coverage](#)
15. [Hyung-In Moon](#) (35) See also: [our coverage](#)
16. [Antonio Orlandi](#) (34) See also: [our coverage](#)
17. [Amelec Viloría aka Jesus Silva](#) (33) See also: [our coverage](#)
18. [Jan Hendrik Schön](#) (32) See also: [our coverage](#)
19. [Dimitris Liakopoulos](#) (31) (NB: We're counting a book he co-authored as a single retraction. The book has 13 retracted chapters with DOIs that are not included in this figure.) See also: [our coverage](#)
20. [Naoki Mori](#) (31) See also: [our coverage](#)
21. [Jose L Calvo-Guirado](#) (30) See also: [our coverage](#)
22. [Soon-Gi Shin](#) (30) See also: [our coverage](#)
23. [Bharat Aggarwal](#) (29) See also: [our coverage](#)
24. [Victor Grech](#) (29) See also: [our coverage](#)
25. [Tao Liu](#) (29) See also: [our coverage](#)
26. [Cheng-Wu Chen](#) (28) See also: [our coverage](#)
27. [A Salar Elahi](#) (27) See also: [our coverage](#)
28. [Prashant K Sharma](#) (27) See also: [our coverage](#)
29. [Richard L E Barnett](#) (26) See also: [our coverage](#)
30. [Dalibor Petkovic](#) (25) See also: [our coverage](#)
31. [Scott Reuben](#) (25) See also: [our coverage](#)



# Perché?



# Valutazione della ricerca

Si basa su indici bibliometrici o,  
per i settori non bibliometrici,  
su liste selezionate di riviste  
(fascia A)





## Cosa stiamo valutando?

- I ricercatori sono valutati sulla base dell'**Impact Factor** delle riviste su cui pubblicano
- **Gli editori commerciali** sono **responsabili** del calcolo dell'Impact Factor delle riviste



# Cosa misura l'Impact Factor?

L'Impact Factor è il numero medio di citazioni per articolo di una rivista, calcolato su un arco temporale di due anni.

$$IF_y = \frac{\text{Citations}_{y-1} + \text{Citations}_{y-2}}{\text{Publications}_{y-1} + \text{Publications}_{y-2}}$$





# Critiche all'uso dell'IF - parte 1

- I ricercatori **all'inizio della carriera** sono penalizzati
- Il contesto della citazione non viene considerato (per esempio le **citazioni negative**)
- Sono influenzati dai limiti dei **database** citazionali (che sono tutti di proprietà di grandi editori scientifici e non sono accessibili)
- Possono essere manipolati dagli autori e dai revisori (autocitazioni e citazioni incrociate)

DORA, 2013, <https://sfdora.org/read/>

McKiernan, et al, 2019. <https://elifesciences.org/articles/47338>

Niles, et al, 2019. <https://www.biorxiv.org/content/10.1101/706622v1>

Alder, et al, 2008. <https://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf>



# Critiche all'uso dell'IF - parte 2

- Non tiene in considerazione il **numero di autori** di un articolo e il loro effettivo contributo
- Non tiene in considerazione aspetti legati alla multidisciplinarietà della ricerca (per esempio le convenzioni citazionali variano a seconda del settore disciplinare)
- Non agevola la libertà scientifica

DORA, 2013, <https://sfdora.org/read/>

McKiernan, et al, 2019. <https://elifesciences.org/articles/47338>

Niles, et al, 2019. <https://www.biorxiv.org/content/10.1101/706622v1>

Alder, et al, 2008. <https://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf>



“Researchers will do anything to publish papers in some journals, including even creating fake authors”

“ [...] publishing papers in certain journals is the only way to earn grants, tenure, and promotions”



**Nonostante tutto, troppi continuano a considerare l'equazione "rivista prestigiosa=ottimo articolo"**

Pensate ad esempio al caso dell'Idrossiclorichina nella ricerca su COVID-19: numerosi articoli pubblicati in riviste scientifiche di alto profilo sono stati ritrattati!



**Ma...buone notizie!**

**Le cose stanno cambiando!**

---



# L'importanza della valutazione

Influenza su tutto il processo scientifico





## Molte le iniziative in corso

**DORA** è un'iniziativa globale che da tempo lavora per trovare nuovi modi di valutazione la ricerca (e in particolare per superare l'IF), che ha raccolto una serie di case studies e partecipa al progetto **TARA** - *Tools to Advance Research Assessment*. Gli enti finanziatori di **CoalitionS** si impegnano a partecipare al rinnovamento dei metodi di valutazione





**Ma la situazione è ancora frammentata**



Il percorso è avviato ma va velocizzato





# La CE si sta muovendo

Cosa succede a livello europeo





# Call for interest

Ora è aperta una call della CE rivolta a organizzazioni interessate a partecipare a una coalizione per cambiare la valutazione

[https://ec.europa.eu/info/news/call-interest-towards-agreement-reforming-research-assessment-2021-dec-16\\_en](https://ec.europa.eu/info/news/call-interest-towards-agreement-reforming-research-assessment-2021-dec-16_en)





# E poi...

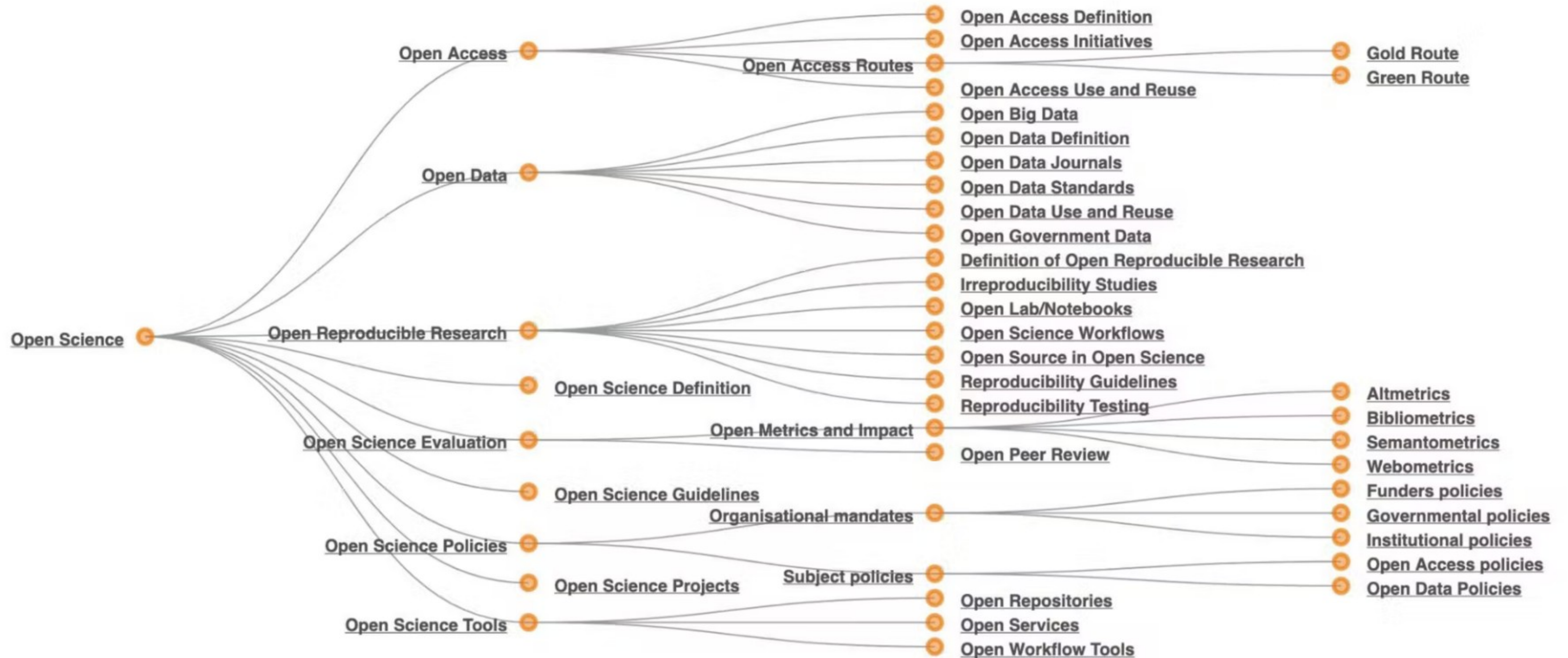
# Open Science!

Che cos'è l'Open Science?

Un sacco di cose



# Open Science è un termine “ombrello”





# Componenti dell'Open Science

Aprire il più possibile e rendere trasparente **ogni passaggio** della ricerca, rivolgendosi ad attori specifici e con mezzi specifici

[Unesco Open Science brochure](#)





# Open Access

I risultati della ricerca sono disponibili per chiunque senza nessuna barriera di accesso





# I vantaggi dell'Open Access

- Maggiore visibilità della ricerca e migliore impatto
- Aiuta a fornire prove dell'impatto
- Migliore reputazione per i ricercatori e le loro istituzioni
- Miglioramento della qualità della ricerca attraverso pratiche di ricerca aperte, trasparenti e riproducibili



CC-BY Danny Kingsley & Sarah Brown



# Integrità

- Trasparenza
- Collaborazione
- Inclusione





# Una scienza aperta alla diversità

- Riconoscere la diversità dei sistemi di conoscenza e delle epistemologie
- Aderire a principi di non discriminazione
- Disponibilità di conoscenze anche per paesi non ricchi





# Non solo pubblicazioni

La science è anche:

- dati
- software
- protocolli
- risultati negativi
- note di laboratorio
- deliverables di progetto
- e molto altro...



# I dati hanno primaria importanza

Controllo  
Validazione  
Follow-up  
Nuovi quesiti di ricerca  
Insegnamento  
...





**Tante attività e prodotti della ricerca non vengono valutati.  
Semplicemente, non si vedono**







Open Science vuol dire collaborazione!

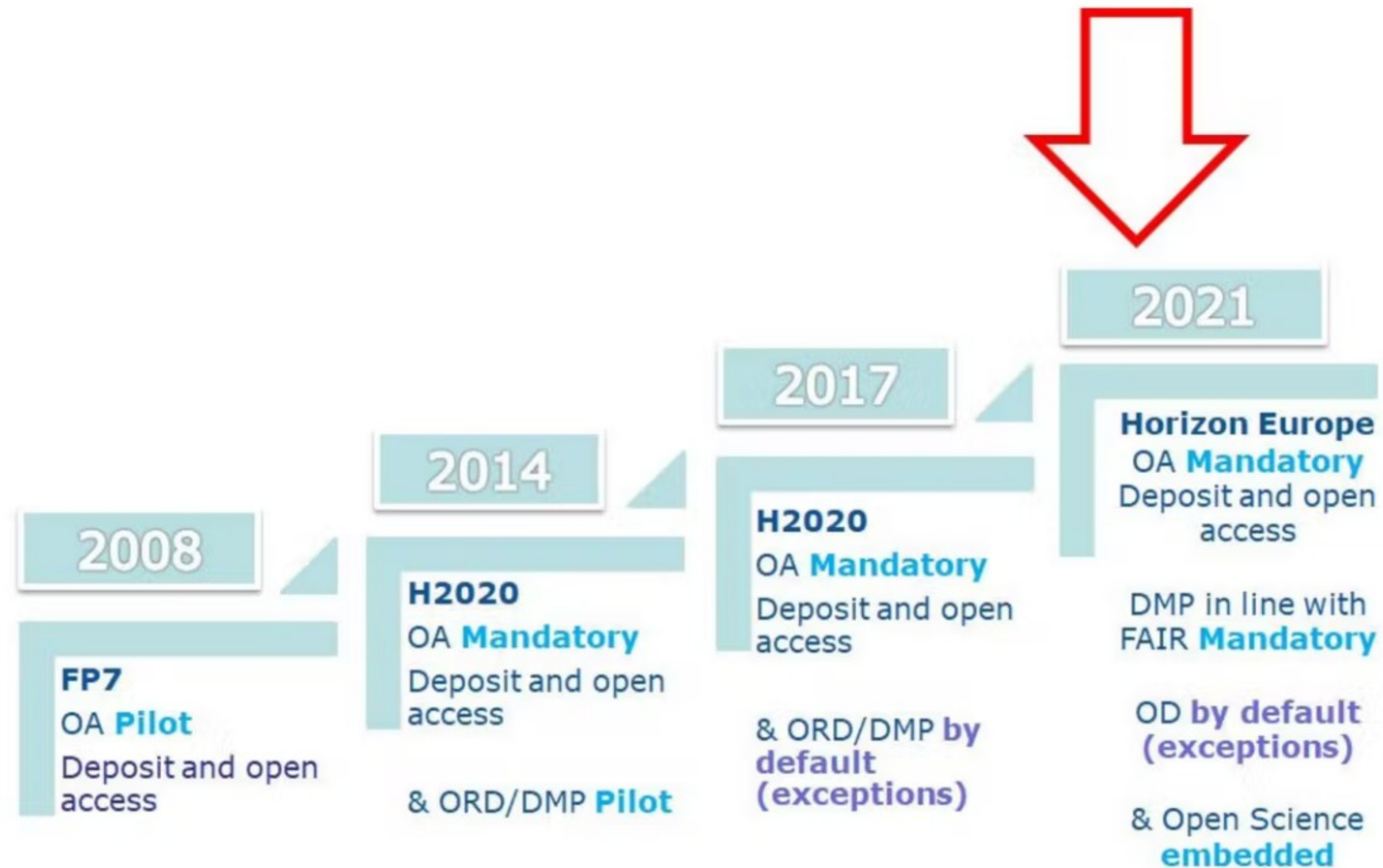


# Open Science as the new normal

Open Science nel programma di finanziamento  
Horizon Europe



# La Commissione europea e la strada verso l'Open Science





# Le principali novità di Horizon Europe in materia di Open Science

- OA obbligatorio per tutte le pubblicazioni derivanti dal progetto
- Deposito della versione OA immediata al momento della pubblicazione (zero embargo)
- Obbligatorio per gli autori mantenere i diritti - right retention strategy
- Costi non rimborsabili per le riviste ibride (ad abbonamento + opzione OA a pagamento per l'autore)





# Le principali novità di Horizon Europe in materia di Open Science

- Fine del programma pilota (nessun opt-out possibile)
- Gestione dei dati secondo i principi FAIR obbligatoria
- DMP obbligatorio per tutti i progetti
- OA è il default, con eccezioni che possono essere motivate nel DMP e accettate dal PO/Reviewers
- Uso di Trusted Repository per il deposito e la conservazione a lungo termine





# Attenzione alla progettazione!

In Horizon Europe la proposta di progetto viene valutata in ottica Open Science





# Open Science nella valutazione della proposta di progetto

**PART-A:** Capacità dei beneficiari di portare a termine gli obiettivi del progetto valutata attraverso

- Achievements: non solo pubblicazioni, ma anche widely-used datasets, software, goods, services or any other achievements relevant to the call content. si deve **fornire una valutazione descrittiva dell'impatto di ciascuno senza usare IF!**
- Projects/activities
- Infrastructures/Equipment

**Part-B:** Methodology

- si dovrà descrivere come le **pratiche di open science** verranno **integrate** nella metodologia
- si dovrà **descrivere la strategia di gestione dei dati/risultati secondo i principi FAIR**

• Describe how appropriate open science practices are implemented as an integral part of the proposed methodology. Show how the choice of practices and their implementation are adapted to the nature of your work, in a way that will increase the chances of the project delivering on its objectives [e.g. 1 page]. If you believe that none of these practices are appropriate for your project, please provide a justification here.

⚠ *Open science is an approach based on open cooperative work and systematic sharing of knowledge and tools as early and widely as possible in the process. Open science practices include early and open sharing of research (for example through preregistration, registered reports, pre-prints, or crowd-sourcing); research output management; measures to ensure reproducibility of research outputs; providing open access to research outputs (such as publications, data, software, models, algorithms, and workflows); participation in open peer-review; and involving all relevant knowledge actors including citizens, civil society and end users in the co-creation of R&I agendas and contents (such as citizen science).*

⚠ *Please note that this question does not refer to communication, dissemination and exploitation described below under 'Impact'.*

• **Research data management and management** generating/collecting data and/or other research output must provide maximum 1 page on how the data/ research principles (Findable, Accessible, Interoperable, Reusable should be specific to your project): [1 page]

**Types of data/research outputs** (e.g. experimental, estimated size; if applicable, combination with, and

**Findability of data/research outputs:** Types of persistent identifiers) and trusted repositories that will be used

**Accessibility of data/research outputs:** IPR considerations and timeline for open access (if open access not provided, explain why); provisions for access to restricted data for verification purposes.

**Interoperability of data/research outputs:** Standards, formats and vocabularies for data and metadata.

**Reusability of data/research outputs:** Licenses for data sharing and re-use (e.g. Creative Commons, Open Data Commons); availability of tools/software/models for data generation and validation/interpretation /re-use.

**Curation and storage/preservation costs;** person/team responsible for data management and quality assurance.

⚠ *Proposals selected for funding under Horizon Europe will need to develop a detailed data management plan (DMP) for making their data/research outputs findable, accessible, interoperable and reusable (FAIR) as a deliverable by month 6 and revised towards the end of a project's lifetime.*

⚠ *For guidance on open science practices and research data management, please refer to the relevant section of the [HE Programme Guide](#) on the Funding & Tenders Portal.*

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]	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).
[Dataset]	
[Software]	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'.
[Good]	
[Service]	
[Other achievement]	



# Evaluation of proposals and Open Science

## “Excellence” criterion

(methodology)

- Up to 1 page to describe OS practices + up to 1 page to describe research data/output management
- Evaluation of the quality of open science practices

**\*\*evaluation concerns mandatory and non-mandatory practices, the latter where appropriate**

**\*\*exception: ERC does not evaluate open science**







## “Quality of implementation” criterion

(capacity of participants and consortium as a whole + list of achievements)

- Explain expertise on OS (if no OS practices are involved then no expertise required)
- List publications, software, data, etc, relevant to the project with qualitative assessment and, where available, persistent identifiers
- Publications expected to be open access; datasets expected to be FAIR and ‘as open as possible, as closed as necessary’. **Significance of publications to be evaluated on the basis of proposers’ qualitative assessment and not per Journal Impact Factor**



# Le pratiche dell'Open Science in Horizon Europe

-  **Condivisione tempestiva e aperta** della ricerca (per esempio pre-prints, pre-registrazione, report registrati eccetera).
-  **Gestione** dei risultati della ricerca - compreso il DMP.
-  Misure per assicurare la **riproducibilità** dei risultati della ricerca.
-  Accesso aperto ai risultati della ricerca - pubblicazioni, dati, software, modelli, algoritmi, flussi di lavoro - attraverso il deposito in archivi affidabili (**trusted repository**).
-  Partecipazione a **peer-review aperta**
-  **Coinvolgimento di tutti gli attori rilevanti della conoscenza**, compresi cittadini, società civile, utilizzatori finali nella co-creazione di contenuti

Per approfondire:

[https://open-science.it/article?rpk=184439&prs\\_sel=p\\_researcher&tpc\\_sel=t\\_regolamenti\\_open\\_science](https://open-science.it/article?rpk=184439&prs_sel=p_researcher&tpc_sel=t_regolamenti_open_science)



# Open Research Europe: la piattaforma di pubblicazione della CE

- Non è una rivista, è una **piattaforma di pubblicazione**.
- Lo scopo è dare ai ricercatori una sede dove pubblicare i risultati delle loro ricerche **finanziate dalla CE**, indipendentemente dal livello percepito di interesse e novità.
- Possono essere pubblicati **studi nulli, risultati negativi o che confermano** le ipotesi della ricerca.
- Lo **scopo della Peer Review** non è di accettare o rigettare un risultato ma di migliorarne la pubblicazione grazie ad uno sforzo di collaborazione tra esperti.
- Il **ruolo dei revisori** è di verificare se la ricerca è solida dal punto di vista della metodologia e se ha merito accademico.
- In linea con i requisiti Open Science di Horizon Europe
- Open Peer Review

The screenshot displays the Open Research Europe interface. At the top, there's a search bar and a 'SUBMIT YOUR RESEARCH' button. The main content area shows a research article titled 'Analysis of the concept of informal economy through 102 definitions: legality or necessity [version 2; peer review: 1 approved, 1 approved with reservations]' by Arturo Luque. The article is marked as 'REVISED'. Below the title, there are tabs for 'Article', 'Authors', and 'Metrics'. The abstract is visible, discussing the complexity of informal economy and the need for a paradigm shift. On the right side, there's a 'Open Peer Review' section showing the reviewer status and reports. A table lists invited reviewers and their reports for different versions of the article.

	1	2
Version 2 (Revision) 10 Jan 22	✓ read	
Version 1 04 Nov 21	↑ read	? read

Invited Reviewers:

- Fahri Fahri Özsungur, Mersin University, Mersin, Turkey
- Edisson Cobo, Universidad Técnica de Ambato, Ambato, Ecuador

Comments on this article: All Comments (0). Sign in to comment.

Sign up for content alerts: Email address \* Sign Up



# Research Data Management

---



## Cosa succede se...

i ricercatori non gestiscono e condividono i dati in modo corretto?





# I dati possono andare persi

## JAMA journal retracts paper when author can't produce original data

In July 2017, a *JAMA* journal called for an investigation into a 2013 paper it had published after concluding that the article had “scientific and ethical concerns.” Now the journal, *JAMA Otolaryngology – Head & Neck Surgery*, is retracting the paper.

The article, “Dexamethasone for the prevention of recurrent laryngeal nerve palsy and other complications after thyroid surgery: a randomized double-blind placebo-controlled trial,” came from a group in Italy led by Mario Schietroma, of the Department of Surgery at the University of L'Aquila, in Abruzzo, Italy. Schietroma, who in December admitted to us that a retracted 2015 paper of his in the *Journal of the American College of Surgeons* suffered from “misinterpretation of the statistical data,” now has four retractions.



*Neither [the original dataset and the approved protocol] have been provided by Dr Schietroma, and the university has informed us that “without those pieces of information the results of the papers under investigation cannot be validated.”*

<https://retractionwatch.com/2018/10/25/jama-journal-retracts-paper-when-author-cant-produce-original-data/>



# L'importanza di depositare i dati della ricerca

MENU ▾

nature

Subscribe



Carlisle has kept going. This year, he warned about dozens of anaesthesia studies by an Italian surgeon, Mario Schietroma at the University of L'Aquila in central Italy, saying that they were not a reliable basis for clinical practice<sup>6</sup>. Myles, who worked on the report with Carlisle, had raised the alarm last year after spotting suspicious similarities in the raw data for control and patient groups in five of Schietroma's papers.



Bottled oxygen, used by anaesthetists during surgery. Credit: Mark Thomas/Alamy

The challenges to Schietroma's claims have had an impact in hospitals around the globe. The World Health Organization (WHO) cited Schietroma's work when, in 2016, it issued a recommendation that anaesthetists should routinely boost the oxygen levels they deliver to patients during and after surgery, to help reduce infection. That was a controversial call: anaesthetists know that in some procedures, too much oxygen can be associated with an increased risk of complications – and the recommendations would have meant hospitals in poorer countries spending more of their budgets on expensive bottled oxygen, Myles says.

The five papers Myles warned about were quickly retracted, and the WHO revised its recommendation from 'strong' to 'conditional', meaning that clinicians have more freedom to make different choices for various patients. Schietroma says his calculations were assessed by an independent statistician and through peer review, and that he purposely selected similar groups of patients, so it's not surprising if the data closely match. He also says he lost



# I dati possono contenere errori

Il caso della teoria sull'austerità.

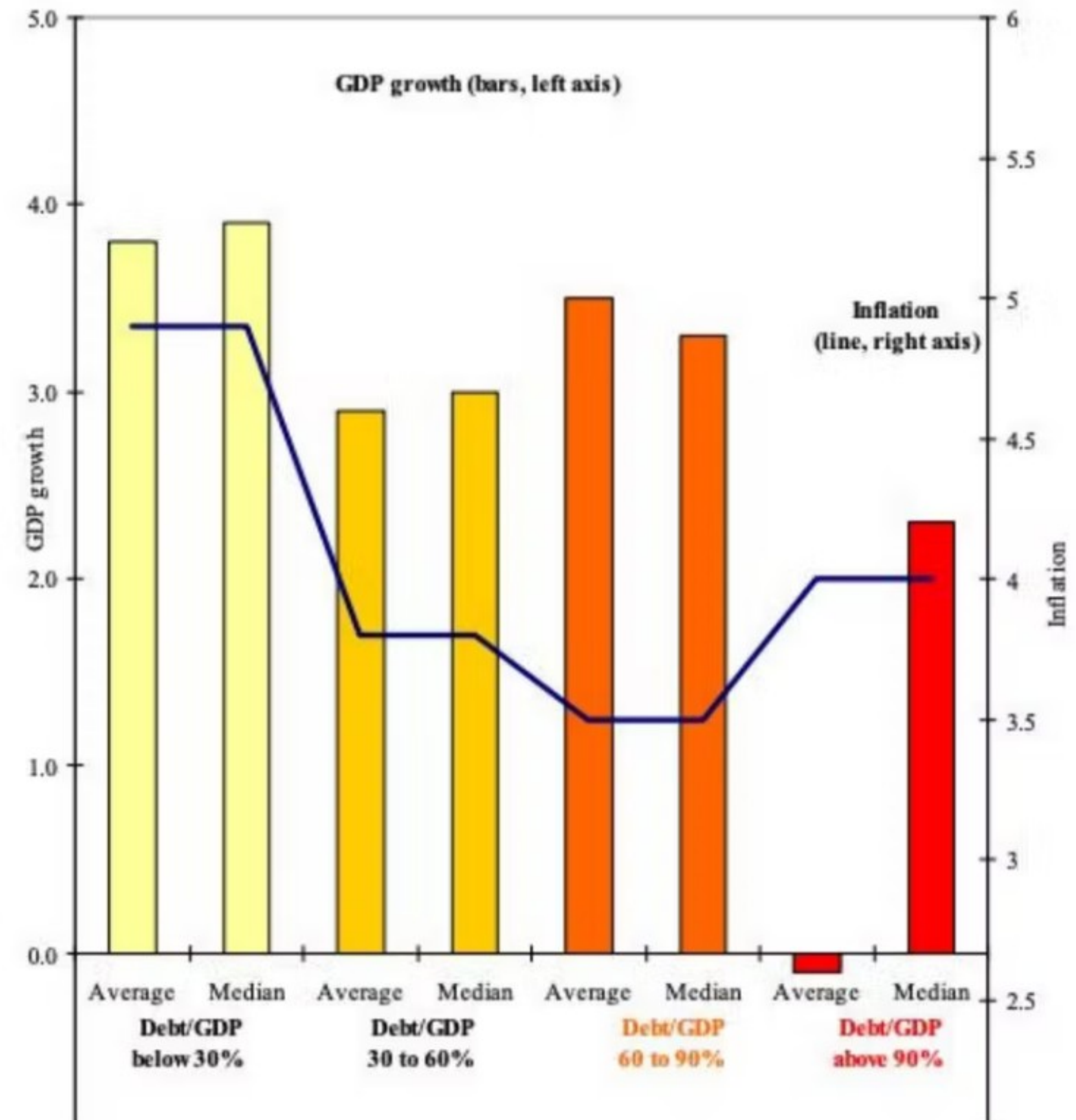
La tesi: la crescita economica rallenta quando la dimensione del debito di un paese supera il 90% del prodotto interno lordo.

I risultati mostrati nell'articolo sono stati usati per sostenere politiche pubbliche di austerità durante la crisi economica.

Ma alcune considerazioni erano basate su calcoli errati.

<https://www.bbc.com/news/magazine-22223190>

Figure 2. Government Debt, Growth, and Inflation: Selected Advanced Economies, 1946-2009





# Errori di calcolo e altri sbagli

The screenshot shows a news article on 'THE CONVERSATION' website. The article title is 'The Reinhart-Rogoff error - or how not to Excel at economics'. The text discusses a critique of a famous 2010 academic paper by Reinhart and Rogoff, which was used by political big-hitters to argue for austerity cuts. The article states that the paper contained significant errors due to the misuse of an Excel spreadsheet. It mentions that Harvard's Carmen Reinhart and Kenneth Rogoff are two of the most respected and influential academic economists active today. The critique was released by Thomas Herndon, Michael Ash, and Robert Pollin at the Political Economy Research Institute at the University of Massachusetts Amherst. The article also notes that Reinhart and Rogoff's work showed average real economic growth slows (a 0.1% decline) when a country's debt rises to more than 90% of gross domestic product (GDP) - and this 90% figure was employed repeatedly in political arguments over high-profile austerity measures.

## Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff

Thomas Herndon\*      Michael Ash      Robert Pollin

April 15, 2013

[Herndon, 2013](#)

JEL CODES: E60, E62, E65

### Abstract

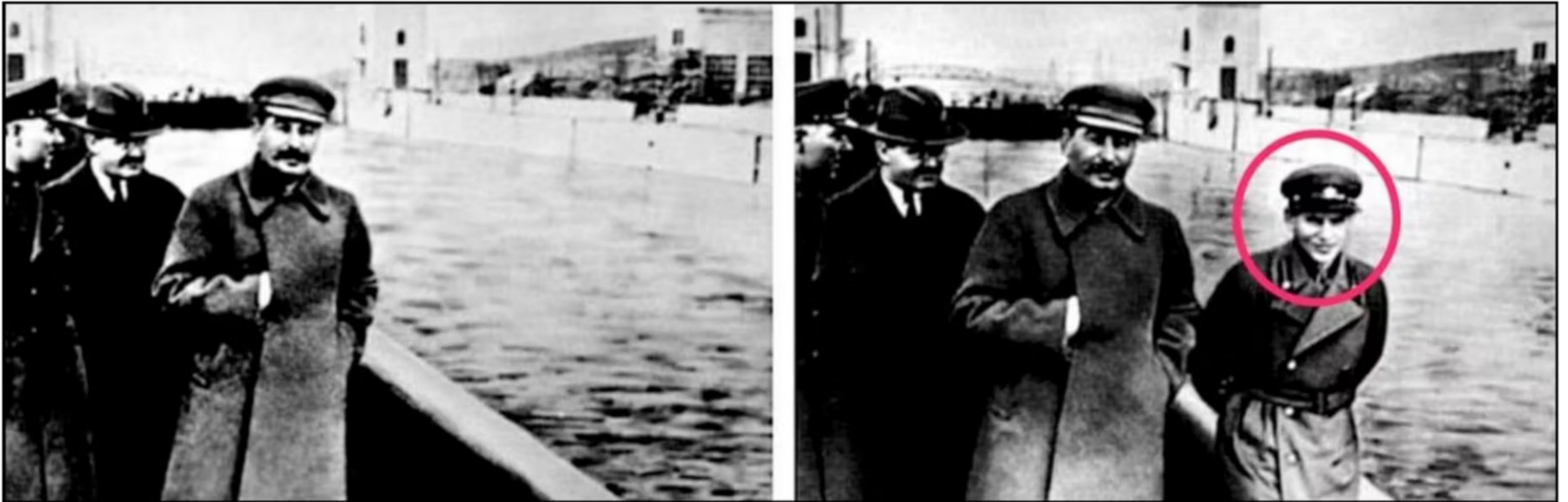
We replicate Reinhart and Rogoff (2010a and 2010b) and find that coding errors, selective exclusion of available data, and unconventional weighting of summary statistics lead to serious errors that inaccurately represent the relationship between public debt and GDP growth among 20 advanced economies in the post-war period. Our finding is that when properly calculated, the average real GDP growth rate for countries carrying a public-debt-to-GDP ratio of over 90 percent is actually 2.2 percent, not -0.1 percent as published in Reinhart and Rogoff. That is, contrary to RR, average GDP growth at public debt/GDP ratios over 90 percent is not dramatically different than when debt/GDP ratios are lower.

We also show how the relationship between public debt and GDP growth varies significantly by time period and country. Overall, the evidence we review contradicts Reinhart and Rogoff's claim to have identified an important stylized fact, that public debt loads greater than 90 percent of GDP consistently reduce GDP growth.



# I dati possono essere manipolati

Nikolai Ivanovich Yezhov era un funzionario della polizia segreta sovietica sotto Joseph Stalin. Dopo aver perso il favore di Stalin, fu ucciso. È stato soprannominato il "commissario scomparso" perché dopo la sua esecuzione, **la sua immagine fu ritoccata** da una foto ufficiale; è uno degli esempi più noti di come la stampa sovietica abbia fatto "scompare" qualcuno che era caduto in disgrazia.



The Newseum (1 September 1999). "The Commissar Vanishes" in The Vanishing Commissar".

Archived from the original on 8 February 2007.

[https://en.wikipedia.org/wiki/Nikolay\\_Yezhov](https://en.wikipedia.org/wiki/Nikolay_Yezhov)



# Quando il problema è l'integrità della ricerca

## REPORT

### Coping with Chaos: How Disordered Contexts Promote Stereotyping and Discrimination

Diederik A. Stapel<sup>1,\*</sup>, Siegwart Lindenberg<sup>1,2,\*</sup>

+ See all authors and affiliations

Science 08 Apr 2011:  
Vol. 332, Issue 6026, pp. 251-253  
DOI: 10.1126/science.1201068

Article

Figures & Data

Info & Metrics

eLetters

PDF

This article has been retracted. Please see:

[Is retracted by - December 02, 2011](#)

#### Abstract

Being the victim of discrimination can have serious negative health- and quality-of-life-related consequences. Yet, could being discriminated against depend on such seemingly trivial matters as garbage on the streets? In this study, we show, in two field experiments, that disordered contexts (such as litter or a broken-up sidewalk and an abandoned bicycle) indeed

**58 articles** published by **Diederik Stapel** were withdrawn because they were based on **invented data**.

His papers had been published in scientific journals considered prestigious (very high IFs!).

Following reports from three doctoral students, the Dutch university for which he worked had started an **investigation**. Stapel then admitted that he had fabricated the data on numerous occasions.

If he had shared his data before, he probably wouldn't have been able to fabricate fakes for so long.

**This case led the Netherlands become one of the pioneer countries in Open Science policy and RDM practices**



# La crisi della riproducibilità

Guardando indietro,  
disponibilità dei dati grezzi alla  
base delle pubblicazioni  
scientifiche diminuisce del 17%  
ogni anno

[https://www.cell.com/current-biology/fulltext/S0960-9822\(13\)01400-0](https://www.cell.com/current-biology/fulltext/S0960-9822(13)01400-0)

Current Biology



Log in Register

Mentimeter

REPORT | VOLUME 24, ISSUE 1, P94-97, JANUARY 06, 2014

PDF (462 KB) Figures

## The Availability of Research Data Declines Rapidly with Article Age

Timothy H. Vines   • Arianne Y.K. Albert • Rose L. Andrew • ... Jean-Sébastien Moore • Sébastien Renault • Diana J. Rennison • Show all authors

Open Archive • Published: December 19, 2013 • DOI: <https://doi.org/10.1016/j.cub.2013.11.014> • Check for updates

Highlights

- We examined the availability of data from 516 studies between 2 and 22 years old
- Policies mandating data archiving at publication are clearly needed

### Summary

Policies ensuring that research data are available on public archives are increasingly being implemented at the government [1], funding agency [2, 3, 4], and journal [5, 6] level. These policies are predicated on the idea that authors are poor stewards of their data, particularly over the long term [7], and indeed many studies have found that authors are often unable or unwilling to share their data [8, 9, 10, 11]. However, there are no systematic estimates of how the availability of research data changes with time since publication. We therefore requested data sets from a relatively homogenous set of 516 articles published between 2 and 22 years ago, and found that availability of the data was strongly affected by article age. For papers where the authors gave the status of their data, the odds of a data set being extant fell by 17% per year. In addition, the odds that we could find a working e-mail address for the first, last, or corresponding author fell by 7% per year. Our results reinforce the notion that, in the long term, research data cannot be reliably preserved by individual researchers, and further demonstrate the urgent need for policies mandating data



26 miliardi di €  
vanno persi **ogni anno** in  
Europa perché i dati  
non vengono gestiti  
correttamente



# Dove sono i dati?

I dati sono la prova alla base di un articolo scientifico, nonostante questo i ricercatori non li condividono e non li gestiscono nel modo corretto.





# I benefici del Research Data Management

## Per i ricercatori

- Maggiore visibilità e opportunità di impatto
- Opportunità di nuove collaborazioni
- Minori rischi di non conformità(aspetti legali, etici, policy istituzionali e degli enti finanziatori)

## Per la scienza

- Facilita la rintracciabilità dei dati e il loro riutilizzo
- Rende possibile nuove ricerche e nuovi approfondimenti attraverso gli stessi dati
- Protezione dei dati
- Supporto all'integrità della ricerca e alla riproducibilità

## Per la società

- Efficienza nell'uso delle risorse pubbliche
- Migliore qualità dei dati a supporto di migliori processi di decision-making
- Opportunità per la citizen science
- Maggiore trasparenza e fiducia nella scienza



**Research Data Management** is simply the effective handling of information that is created in the course of research.

[How and why you should manage your research data: a guide for researchers](#)

[An introduction to engaging with research data management processes.](#)

[Caroline Ingram, JISC Guides](#)









“Research data management (RDM) is the **careful handling and organization** of research data during the entire research cycle, with the aim of making the research process as **efficient** as possible and to facilitate cooperation with others. More specifically, RDM helps to **protect data**, it facilitates in **sharing** the data with others and it ensures that research data is findable, accessible and (re)usable”

Smits, D.A.B. and Teperek, M., 2020. Research Data Management for Master's Students: From Awareness to Action. *Data Science Journal*, 19(1), p.30. DOI: <http://doi.org/10.5334/dsj-2020-030>



# Perché dovresti prenderti cura dei dati?

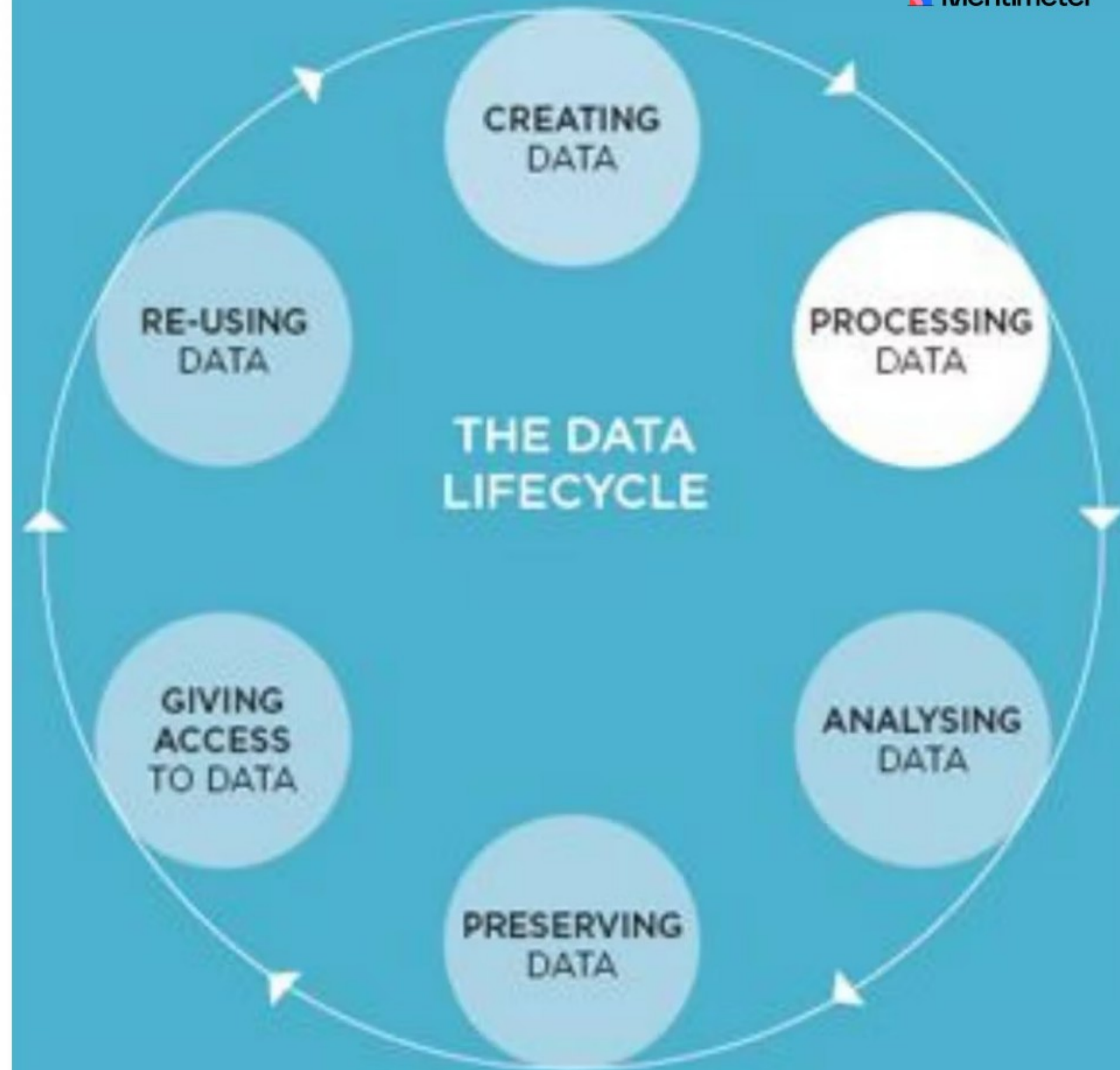
-  Per non perderli
-  Organizzare i tuoi dati serve a lavorare meglio e con maggiore efficienza
-  Alcuni dati sono **unici e irriproducibili** (per esempio rilevazioni meteorologiche, dati da osservazioni sul campo ecc) e a maggior ragione bisogna trattarli in modo molto accurato
-  Per migliorare l'**integrità della ricerca**
-  Per consentire processi di **validazione e controllo**
-  Perché qualcun altro potrebbe riusare quei dati in futuro per altri importanti lavori





# Le pratiche RDM comprendono attività in ogni fase del lavoro con i dati

Prima, durante e dopo il progetto di ricerca. Le scelte fatte in ciascuna fase influenzano quella successiva.





# RDM

Research Data Management

## Planning

Workflow design  
Legal and ethical aspects  
DMP  
Check list  
Conventions  
Policy framework  
Mandates

## Data COLLECTION

File formats  
File naming  
Folder structure  
Documentation  
Storage and back-up  
Collaboration  
Standards/Best practices

## Data PRESERVATION

Persistent Identifiers  
Metadata  
Trusted Repository  
Data curation  
Data Selection

## Data SHARING

Access rights  
Licences  
Software sharing  
Reuse  
Policies



# **FAIR principles**

---



# The FAIR Principles



## Findable

Other can find your data



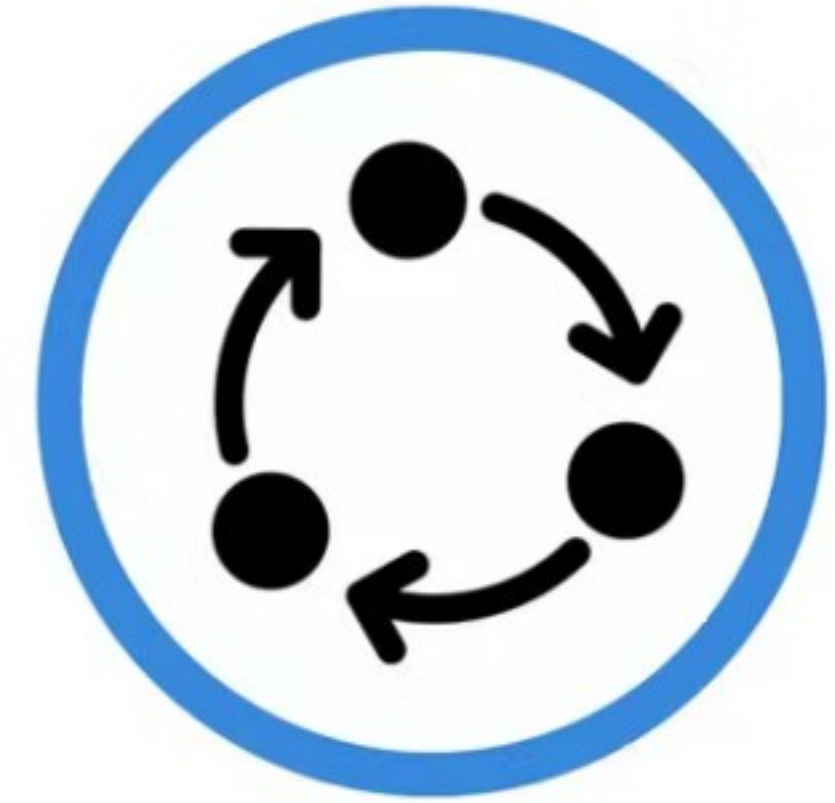
## Accessible

Your data is accessible to others



## Interoperable

Your data can be integrated with other data and/or they can be easily used and read by machines.



## Reusable

Your data can be reused by others in new research

These slides are inspired by Martínez-Lavanchy, P.M., Hüser, F.J., Buss, M.C.H., Andersen, J.J., Begtrup, J.W. (2019). 'FAIR Principles'. In: Holmstrand, K.F., den Boer, S.P.A., Vlachos, E., Martínez-Lavanchy, P.M., Hansen, K.K. (Eds.), Research Data Management (eLearning course). doi: 10.11581 / dtu: 00000049

Link to the original video: <https://vidensportal.deic.dk/RDMelearn>



# I principi FAIR

- FAIR indica una lista di principi che possono aiutarvi a rendere i vostri dati pronti per l'Open Science
- Sono principi, non standard!
- Sono stati messi a punto per consentire un uso ottimale dei dati e dei metodi di ricerca
- Un gruppo di esperti ha progettato i principi FAIR tra il 2014 e il 2016
- Si tratta di 15 principi

www.nature.com/scientificdata

# SCIENTIFIC DATA

Amended: Addendum

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

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

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

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measurable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

**Supporting discovery through good data management**  
Good data management is not a goal in itself, but rather is the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse by the community after the data publication process. Unfortunately, the existing digital ecosystem surrounding scholarly data publication prevents us from extracting maximum benefit from our



# Applicando i principi FAIR

- Produrrà dati di elevata qualità
- Potrai aumentare l'impatto della tua ricerca
- Potrai avere maggiore riconoscimento all'interno e oltre la tua comunità di ricerca
- Sarete conformi agli obblighi derivanti da alcuni enti finanziatori (per esempio quelli della Comunità Europea)

L'applicazione dei principi FAIR dipende fortemente dalla disciplina specifica e dal modo in cui lavora il singolo ricercatore


**Non esiste una ricetta sempre valida!**



# La strategia CISUP

---





La gestione dei risultati della ricerca secondo i principi FAIR necessita di una adeguata strategia istituzionale a supporto del ricercatore



# Il gruppo di lavoro

Il gruppo si occuperà, in contatto diretto con ricercatori, tecnici e utenti, e con il supporto del Centro di Competenza ICDI per l'Open Science, di:

- definire la **strategia** per la gestione dei dati CISUP (formazione, supporto, ...)
- scrivere una **policy**, al pari di altri enti e laboratori europei
- mettere a disposizione dei ricercatori e dei tecnici CISUP gli **strumenti** necessari per la gestione dei dati (flussi di lavoro, strumenti di collaborazione e condivisione, ...)

## Composizione:

- una dozzina di persone, suddivise in 3-4 sottogruppi tematici
- priorità ai ricercatori quotidianamente impegnati nei lab CISUP e a quelli dipartimentali coinvolti
- **autocandidature aperte!**





# Discussione

è utile avere tutti i dati prodotti organizzati e salvati. Sono il risultato del lavoro dei ricercatori UNIFI e "patrimonio" dell'Ateneo

Se un repository è su EOSC, posso avere la certezza che questo sia trusted?

Quando un repository dei dati è "trusted"?

Depositare i dati delle ricerche UNIFI nel data center di ateneo e renderli interoperabili e completamente accessibili potrebbe essere una soluzione?

Potremmo creare un repository trusted qua a Pisa?

Il re-use dei dati per altri studi vede il riconoscimento del ricercatore/Ateneo che li ha originati? Se sì in quale forma?