



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

# OpenCitations

La valutazione della ricerca alla  
prova dei fatti, 17 gennaio 2024

## Dati bibliografici e citazionali aperti per la trasparenza della valutazione della ricerca: il caso di OpenCitations

Silvio Peroni

Research Centre for Open Scholarly Metadata, Dipartimento di Filologia Classica e  
Italianistica, Università di Bologna, Bologna, Italia

[silvio.peroni@unibo.it](mailto:silvio.peroni@unibo.it) – [@essepuntato](https://www.essepuntato.it) – <https://www.unibo.it/sitoweb/silvio.peroni/en>

OpenCitations, Direttore

[silvio.peroni@opencitations.net](mailto:silvio.peroni@opencitations.net) – [@opencitations](https://www.opencitations.net) – <https://opencitations.net>



RESEARCH CENTRE  
FOR OPEN SCHOLARLY METADATA



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
DEPARTMENT OF CLASSICAL PHILOLOGY  
AND ITALIAN STUDIES



# Un ringraziamento

## Direttori

David  
Shotton



Silvio  
Peroni

## Sviluppo

Arcangelo  
Massari



Arianna  
Moretti

Elia Rizzetto



Ivan  
Heibi

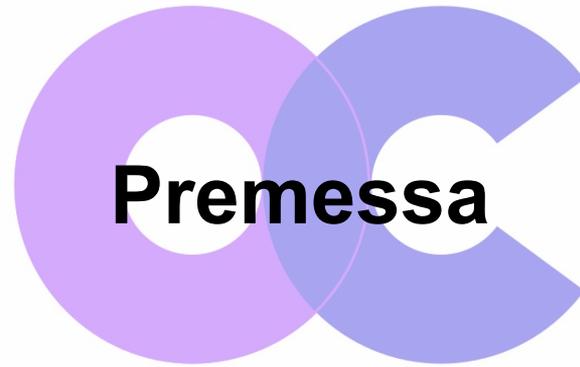
## Amministrazione e comunicazione

Claudio  
Fabbri



Chiara  
Di Giambattista

... e altri fantastici contributi di (in ordine alfabetico) Davide Brembilla, Fabio Mariani, Gabriele Pisciotta, Giuseppe Grieco, Lorenzo Paolini, Marilena Daquino, Marta Soricetti, Olga Pagnotta, Philipp Mayr, Sahar Vahdati, Simone Persiani



**Premessa**

# Che cos'è una citazione aperta?

Una **citazione bibliografica** è un **link concettuale** da un'entità citante a un'entità citata creato con il proposito di dare credito al contributo presentato, dai relativi autori, nell'entità citata

Questo link è definito includendo un riferimento bibliografico nella lista dei riferimenti bibliografici dell'entità citante e denotando questo riferimento bibliografico attraverso un puntatore testuale per il riferimento bibliografico a cui si vuol fare riferimento (ad esempio “[3]” o “(Doe et al., 2021)”)

I **dati** che descrivono una specifica citazione devono includere:

- la **rappresentazione** di questo link concettuale
- i **metadati di base** dell'entità citante e dell'entità citata, ovvero dell'informazione che sia sufficiente per creare il riferimento bibliografico testuale di ogni entità coinvolta

Una citazione bibliografica è una **citazione aperta** quando i dati che la definiscono sono conformi ai seguenti cinque principi

strutturati

```
"reference": [{
  "issue": "2",
  "key": "10.7717/peerj.4375/ref-11",
  "doi-asserted-by": "crossref",
  "first-page": "237",
  "DOI": "10.1002/asi.22963",
  "article-title": "Anatomy of green open access",
  "volume": "65",
  "author": "Björk",
  "year": "2014",
  "journal-title": "Journal of the Association for
},
...
```

separati



aperti

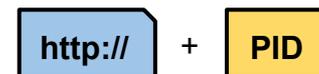


without  
restrictions to  
maximise reuse

identificabili



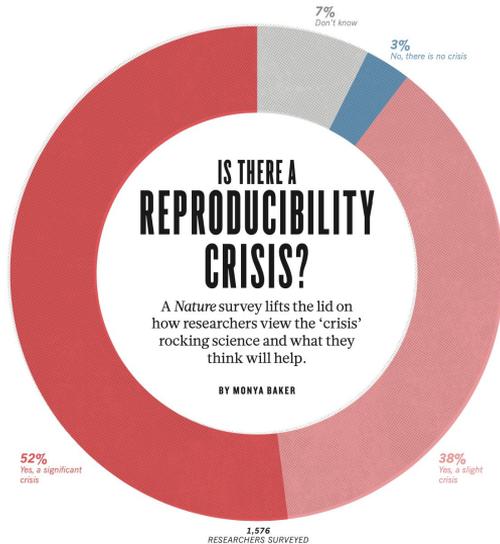
disponibili



# Riproducibilità e valutazione della ricerca

Riproducibilità: “whatever results we compute and report, **someone else** who uses exactly the **same data, methods or algorithms** should be **able to reproduce the same results**”

(Meng, X.-L., 2020)



>70% hanno provato e fallito a riprodurre un esperimento fatto da terzi

>50% hanno fallito nel riprodurre il proprio esperimento

(Baker, M., 2016)



Questa “crisi” relativa alla riproducibilità riguarda sia la ricerca, sia i metodi usati per valutarla

National Plan For Open Science,  
<https://www.ouvri.lascience.fr/national-plan-for-open-science-4th-july-2018/>



San Francisco Declaration on Research  
Assessment, <https://sfdora.org>



Leiden Manifesto for Research Metrics,  
<http://www.leidenmanifesto.org/>



Towards a reform of the research  
assessment system: scoping report,  
<https://doi.org/10.2777/707440>



“Assessment system for researchers and research institutions must be updated to **reflect the principles and practices of open science**”

“**Be open and transparent** by providing data and methods used to calculate all metrics”

“Provide the data under a license that **allows unrestricted reuse**, and provide **computational access** to data”

“Allow those evaluated to **verify data and analysis**”

“Keep data collection and analytical processes **open, transparent and simple**”

“Ensure **independence and transparency** of the data, infrastructure and criteria necessary for research assessment and for determining research impacts [...] by ensuring **control and ownership by the research community over critical infrastructures and tools**”

# Riforma della valutazione della ricerca



<https://coara.eu/>

## 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.

## UNESCO Recommendation on Open Science

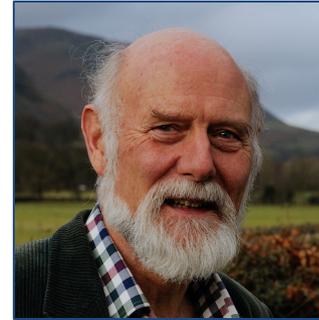
<https://unesdoc.unesco.org/ark:/48223/pf0000379949>

**Open science infrastructures** refer to shared research infrastructures (virtual or physical, including major scientific equipment or sets of instruments, knowledge-based resources such as collections, journals and open access publication platforms, repositories, archives and scientific data, current research information systems, **open bibliometrics and scientometrics systems for assessing and analysing scientific domains**, open computational and data manipulation service infrastructures that enable collaborative and multidisciplinary data analysis and digital infrastructures) that are needed to support open science and serve the needs of different communities.

Dati citazionali chiusi e a pagamento non sono conformi con i criteri di trasparenza, replicabilità e verificabilità richiesti

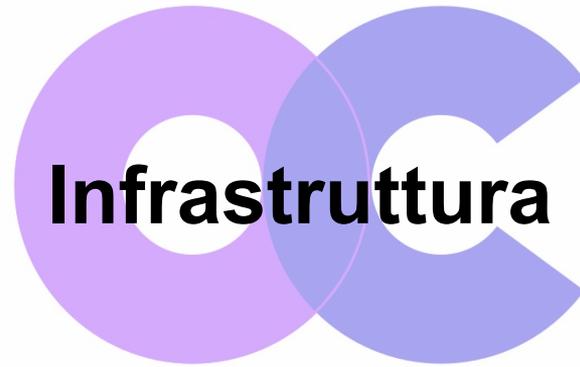
# Verso la trasparenza dei dati citazionali

*“In this open-access age, **it is a scandal** that reference lists from journal articles — core elements of scholarly communication that permit the attribution of credit and integrate our independent research endeavours — are not readily and freely available for use by all scholars.”*



La libera circolazione di dati bibliografici e citazionali aperti è un requisito cruciale per permettere la **creazione di metriche riproducibili da usare negli esercizi di valutazione della ricerca**

Fin dai suoi albori nel 2010, **OpenCitations** è stato creato con l'idea di fornire libero accesso a dati bibliografici e citazionali globali



**Infrastruttura**

# OpenCitations: missione e servizi

OpenCitations è un’infrastruttura aperta guidata dalla comunità che ha come **missione** quella di **raccogliere** e **pubblicare in modo aperto** i metadati descrittivi delle pubblicazioni accademiche e le relative citazioni che le interconnettono, oltre che di **preservare** e permettere l’accesso a questa informazioni archiviando i relativi dati in opportuni depositi sicuri

OpenCitations è gestita dal [Research Centre for Open Scholarly Metadata](#), un centro di ricerca indipendente dell’[Università di Bologna](#)

OpenCitations fornisce

- data model: l’[OpenCitations Data Model](#) (OCDM, basato sulle [SPAR Ontologies](#))
- database bibliografici e citazionali: [OpenCitations Meta](#) and [OpenCitations Index](#)
- software: [GitHub repository](#) rilasciato con licenze open source
- servizi online: [dumps](#), [REST APIs](#), [SPARQL endpoints](#)



# Le nostre sorgenti

OpenCitations Index è un indice bibliografico di citazioni tra pubblicazioni (di varia natura) creato a partire da diverse sorgenti bibliografiche, in cui le citazioni sono trattate come entità di primo livello, caratterizzate da metadati descrittivi e identificate da un Open Citation Identifier (OCI, <https://identifiers.org/oci>)

OpenCitations Meta è un database che conserva e mette a disposizione metadati bibliografici (titolo, tipo, sede editoriale, volume, fascicolo, pagine, data di pubblicazione, identificativi, autori, curatori, editori) per tutte le pubblicazioni coinvolte nell'OpenCitations Index – ogni entità in OpenCitations Meta è associata ad un identificativo interno persistente chiamato OpenCitations Meta Identifier (OMID, <https://identifiers.org/omid>)

	INDEX	Crossref	DataCite	PubMed	OpenAire	JaLC
INDEX	1,975,552,846	1,563,218,160	169,814,412	695,988,810	14,645,838	396,788
Crossref		1,100,963,346	27,051	458,309,297	3,917,329	1,137
DataCite			169,663,255	9,623	114,483	0
PubMed				237,208,867	9,711,789	125
OpenAire					1,067,712	0
JaLC						395,526

**1,97 miliardi di link citazionali**



Crossref  
DOI



DataCite  
DOI



NIH-OCC  
PMID

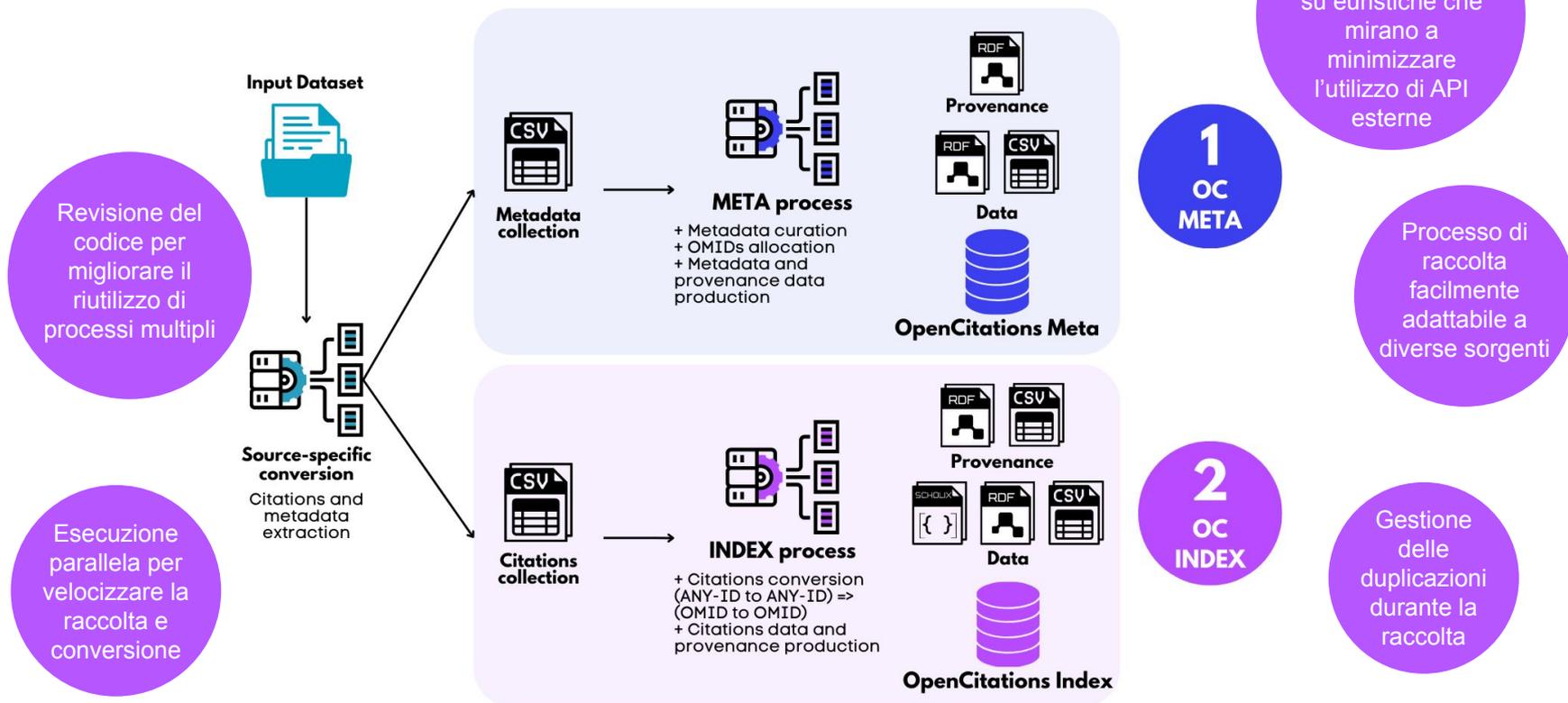


OpenAIRE  
DOI, PMID,  
PMC, ARXIV



JaLC  
DOI

# Processo di raccolta dati



# Provenance e tracciamento dei cambiamenti

-  dati aggiunti
-  dati rimossi
-  dati stabili
-  delta:  
dati aggiunti e rimossi

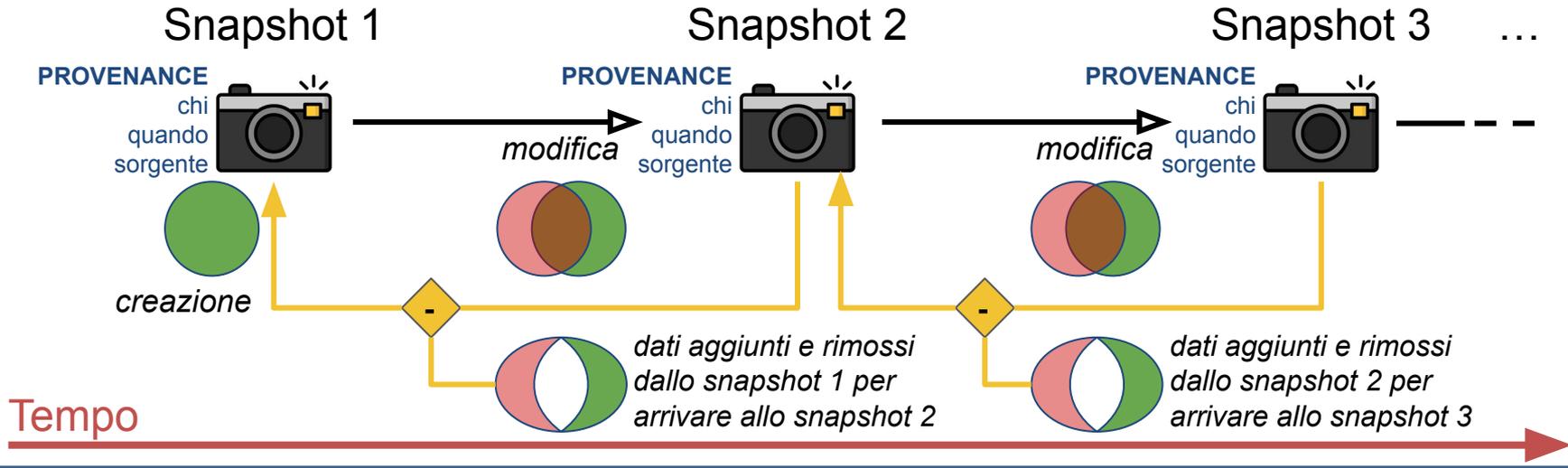


snapshot: la fotografia dei dati associati ad un'entità in un certo tempo



Un'entità  
ad esempio un articolo di rivista

Per tenere traccia della storia di un'entità, dobbiamo salvare i dati del suo ultimo snapshot più tutti i delta costruiti per modificare gli snapshot precedenti



# Modificare i dati

Problema: permettere ad esperti di dominio (ad esempio bibliotecari) di modificare i dati di OpenCitations in modo sicuro, tenendo traccia della loro provenance e dei loro cambiamenti nel tempo

## Soluzione

- Interfacce Web create SHACL document (che definisce le tipologie di entità e le loro relative proprietà relative) + un file JSON per ulteriori parametri di configurazione
- Tracciamento automatico di informazioni di provenance e dei cambiamenti

About <https://w3id.org/oc/meta/br/06014>

Type  
Bibliographic Resource

journal article

Identifier  
[doi:10.1002/\(sici\)1097-0142\(19990501\)85:9<2017::aid-cnrc20>3.0.co;2-v](https://doi.org/10.1002/(sici)1097-0142(19990501)85:9<2017::aid-cnrc20>3.0.co;2-v)

Title  
Allelic Imbalance In The Clonal Evolution Of Prostate Carcinoma

Author  
Cheng, Liang  
Bostwick, David G.

1:44:48 PM NOVEMBER 27, 2023 – 2:44:55 PM NOVEMBER 27, 2023

### SNAPSHOT 2

Responsible agent: <https://orcid.org/0000-0002-8420-0696>  
Primary source: Unknown  
Description: The entity <https://w3id.org/oc/meta/br/06014> was modified.  
Deletions: Publication Date 1999-05-01  
Additions: Publication Date 1999-06-01

[View version](#)

Timeline: NOV.11, NOV.13, NOV.15, NOV.17, NOV.19, NOV.21, NOV.23, NOV.25, NOV.27, NOV.29, DEC.1, DEC.3, DEC.4, DEC.6, DEC.8, DEC.10, DEC.12, DEC.14

# OpenCitations: comunità

OpenCitations collabora con una rete estesa di infrastrutture Open Science e organizzazioni, ed è coinvolta in progetti che si focalizzano sulla creazione di un ambiente di ricerca aperto ed egualitario

DIRECTORY OF OPEN ACCESS BOOKS  
OPEN ACCESS PUBLISHING FOR EUROPEAN NETWORKS  
PUBLIC KNOWLEDGE PROJECT  
OPENCITATIONS

## SCOSS LAUNCHES SECOND FUNDING CYCLE

Read about these essential services, their funding goals and how your institution can help support them at [www.scoss.org](http://www.scoss.org).



RESEARCH INFRASTRUCTURE FOR SCIENCE AND INNOVATION POLICY STUDIES

OpenCitations è inclusa nell'[EOSC Marketplace](#) e nel [JISC Catalogue](#);

OpenCitations è membro della "[SCOSS Family](#)" del gruppo sui "[POSI Adopters](#)"

# Un'infrastruttura basata e guidata dalla comunità



## COINVOLGIMENTO DELLA COMUNITÀ

Questo coinvolgimento è declinato in tre aspetti:

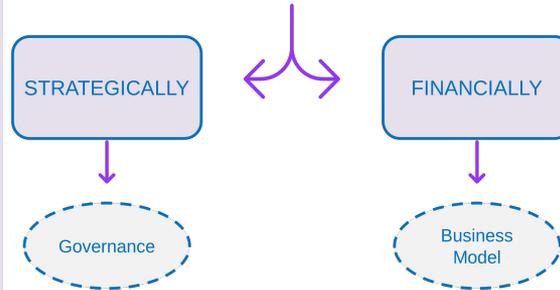
**Partnership collaborative** con individui, infrastrutture, servizi e data provider, che includono il coinvolgimento di parti terze (ad esempio bibliotecari universitari) per fornire e curare i dati di OpenCitations.

**Crowdfunding di comunità** per fornire supporto finanziario a OpenCitations.

**Partecipazione della comunità** nello sviluppo e nella gestione di OpenCitations.



## STAKEHOLDERS INTEGRATI IN OPENCITATIONS



## LAVORARE SUI PUNTI DEBOLI

Per guidare gli sviluppi non tecnici di OpenCitations, i membri della OpenCitations International Advisory Board sono coinvolti in **due gruppi di lavoro** – focalizzati sui temi riguardanti l'evoluzione dei processi di gestione dell'infrastruttura (**Governance Evolution**) e nella creazione della rete di comunità (**Community Building**) – che si incontrano diverse volte durante l'anno e restituiscono le loro raccomandazioni ai Direttori ed all'Advisory Board.



**Grazie per  
l'attenzione**