

Open Science an introduction

Pascal Aventurier (IST/MCST IRD)

OLInFER'2018. UCI. La Habana. November 12th, 2018

 @IST_IRD email : pascal.aventurier@ird.fr

<http://horizon.documentation.ird.fr>

https://www.netvibes.com/doc_ird_de_montpellier#aides_a_la_publication

Open Science



About

Resources

Events

Courses

News

Search for...

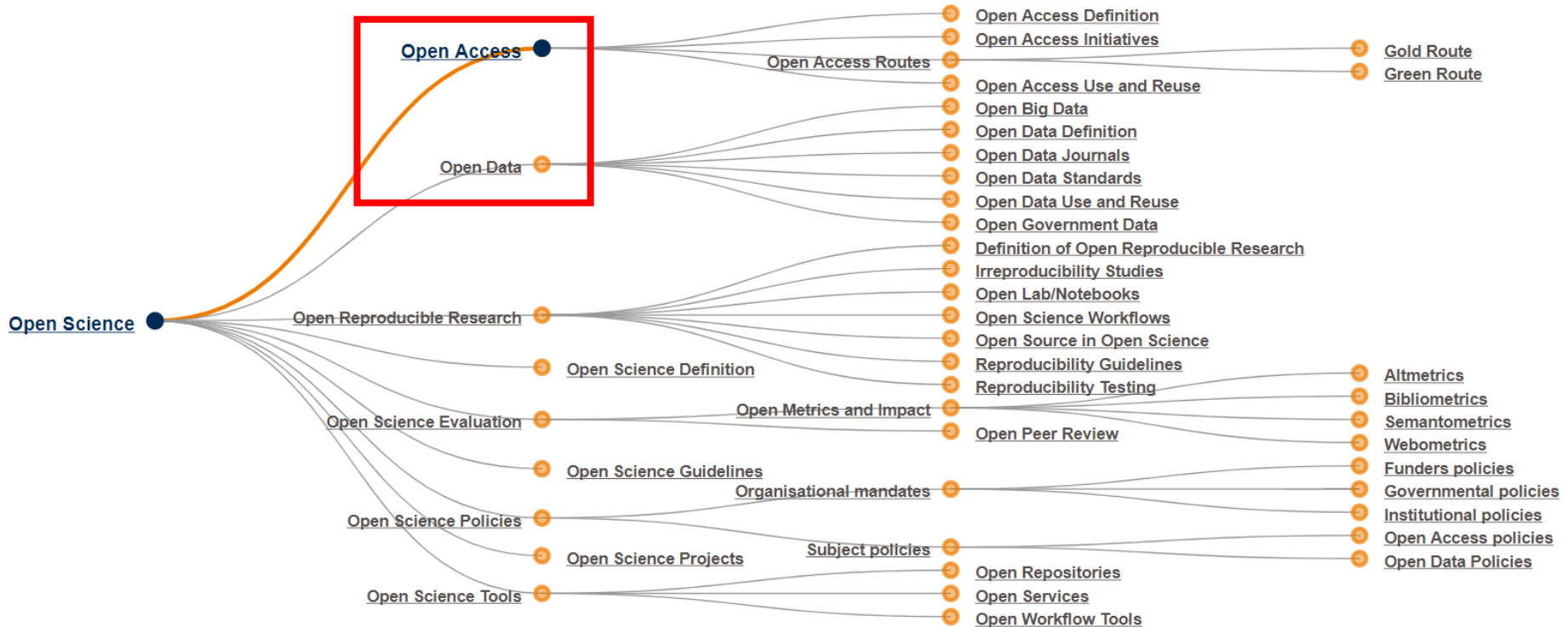


Open Science is the practice of science in such a way that others can collaborate and contribute, where **research data, lab notes and other research processes are freely available**, under terms that enable **reuse, redistribution and reproduction** of the research and its underlying data and methods.

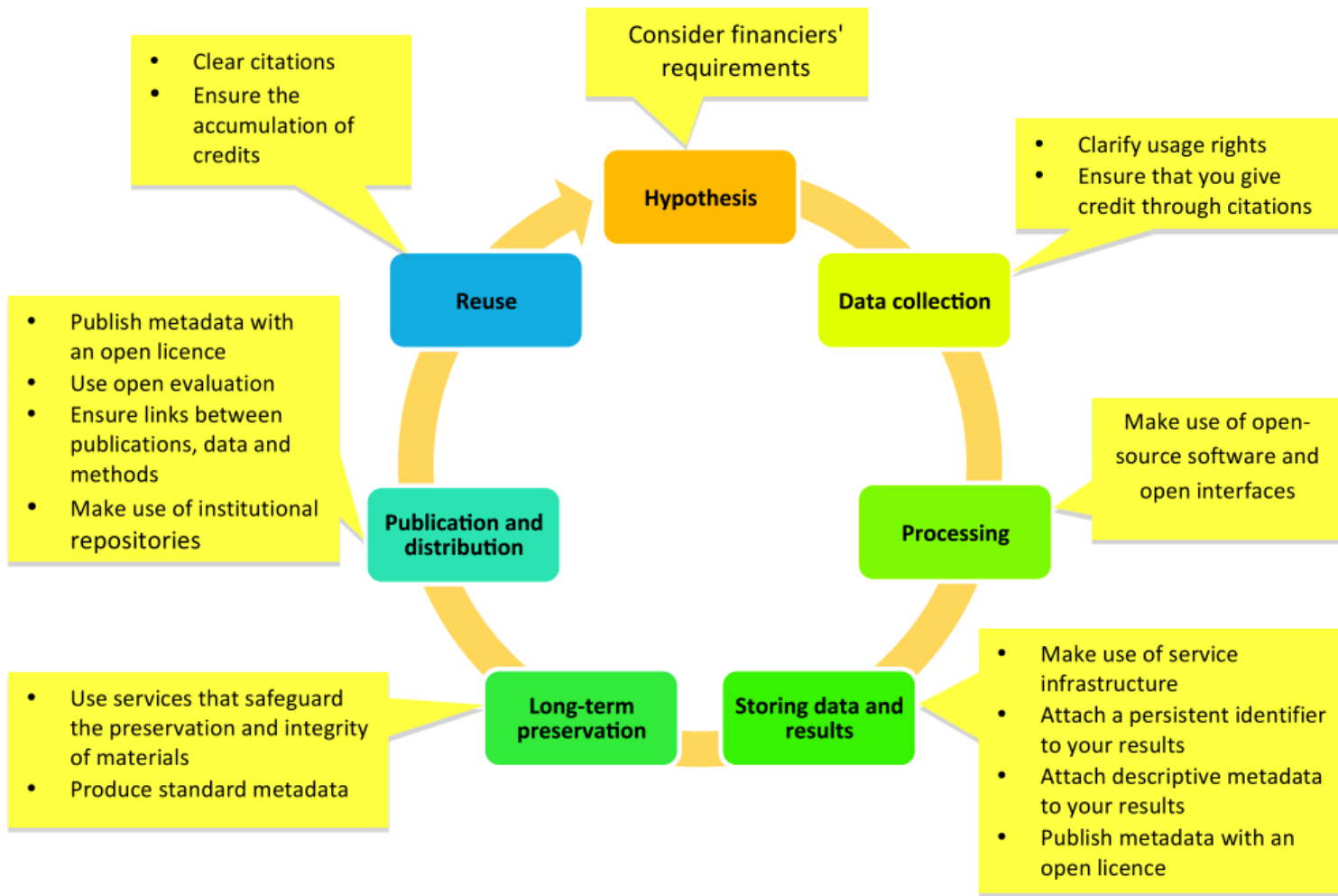
<https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>



Open Science Taxonomy



<https://www.fosteropenscience.eu/taxonomy/>



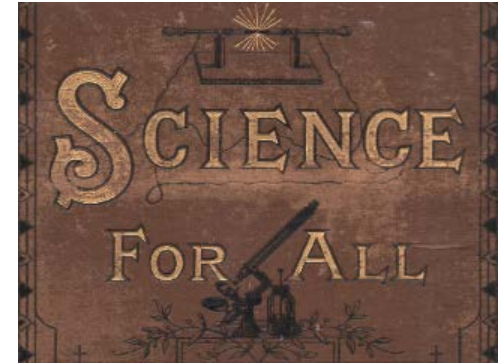
<https://www.fosteropenscience.eu/content/what-open-science-introduction>

The Benefits of Open Science



BETTER POLICIES FOR BETTER LIVES

<http://www.oecd.org/sti/inno/open-science.htm>



- ◆ Promotes a more accurate verification of scientific results.
- ◆ Reduces duplication in collecting, creating, transferring and re-using scientific material.
- ◆ Increases productivity in an era of tight budgets.
- ◆ Results in great innovation potential and increased consumer choice from public research.
- ◆ Promotes citizens' trust in science. Greater citizen engagement leads to active participation in scientific experiments and data collection.

Open science : an opportunity for the Global South

Open innovation platform for drug discovery

The European Lead Factory is a novel, open-innovation platform for drug discovery, managed by an international consortium of 30 partners and funded under the European Innovative Medicines Initiative (IMI). It has been designed to offer unrivalled opportunities to non-contractual public and private parties for the discovery of new drug lead molecules. European academics and SMEs benefit from access to an 'industry-like' discovery platform that encourages active participation. The ultimate goal is to address neglected diseases including bacterial infections, psychiatric disorders and oncology.

- Drug Discovery
- 30 partners
- Funded under the European Innovative Medicine Initiative
- Discovery of new drug lead molecule
- The platform encourages active participation
- Neglected diseases



<https://www.fosteropenscience.eu/learning/open-science-id-innovation/#/id/5aaa7deddd1827131b90e028>
New Pharmaceutical Commons: Transforming Drug discovery 2017
<https://www.oxfordmartin.ox.ac.uk/publications/view/2613>



**Amsterdam Call for Action
on Open Science**

- 1-Full Open Access to all publications**
- 2-A fundamentally new approach towards optimal reuse of research Data**
- 3- New assessment, reward and evaluation systems**
- 4-Alignment of policies and exchange of best practices**

RU
2016

Digital Transformation Agency sponsors hackathon to unlock potential benefits of open data in Australia

BY:  **TERESA UMALI** GEOGRAPHY: **AUSTRALIA** PUBLISHED: **4 AUG 2018**

[SUBSCRIBE TO OUR NEWSLETTER](#)



LINKEDIN



FACEBOOK



TWITTER



EMAIL

Open data holds a potential wealth of information. A recent research suggested that making data publicly available for re-use could generate up to A\$ 25 billion each year. GovHack 2018 can help find possible uses for it, eventually discovering its potential value.

Teams of developers, coders, designers, analysts and others will participate in the [GovHack 2018](#) event. In the open data hackathon, the teams will be using open data to develop solutions and services that are aimed at benefiting the broader community.

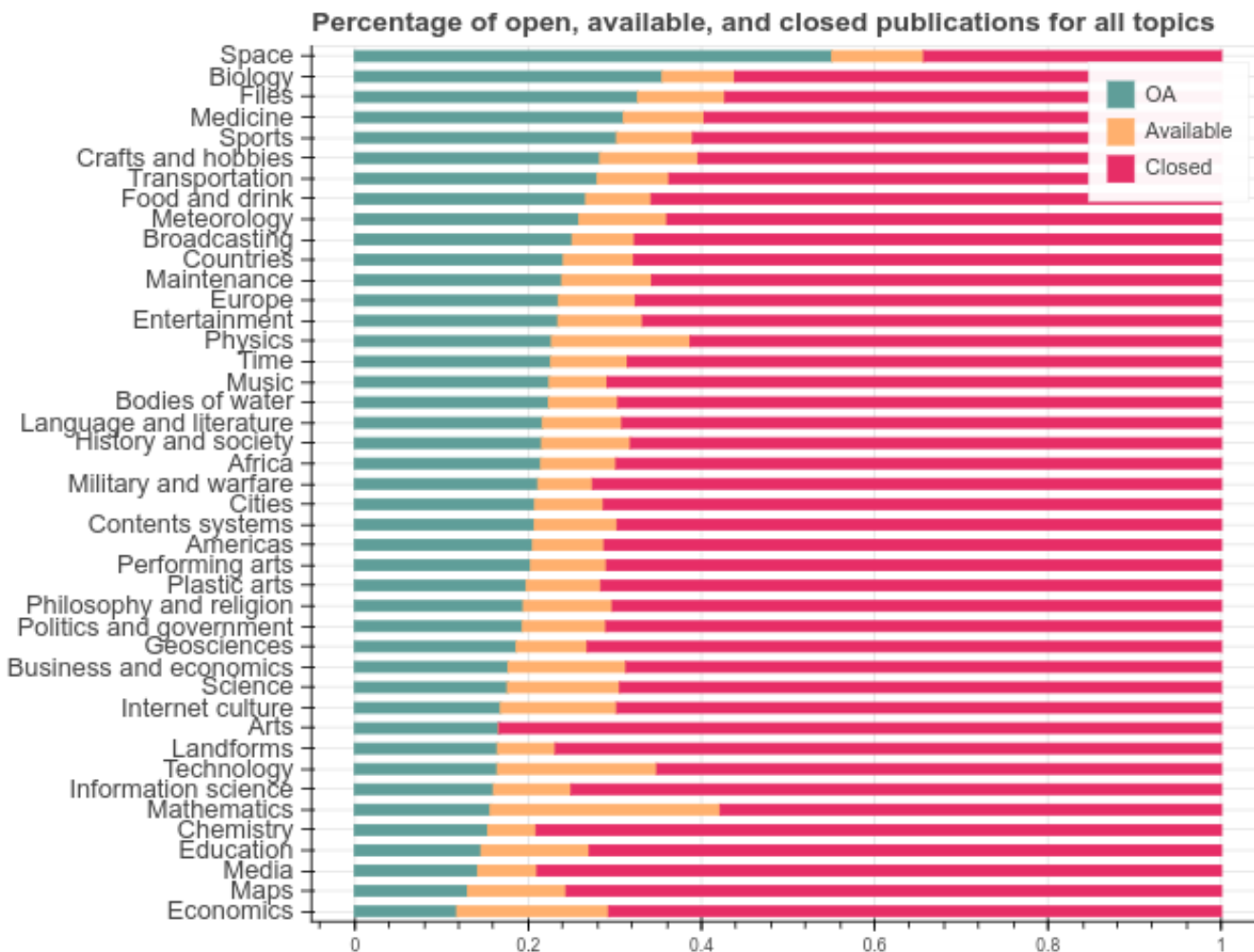
Commercial publishers

Elsevier, Springer + 35% benefit

WHAT IS THE PROBLEM?



How Many Wikipedia References Are Available to Read?



« How Many Wikipedia References Are Available to Read? We Measured the Proportion of Open Access Sources across Languages and Topics. » 2018. *Wikimedia Foundation* (blog). 20 août 2018. <https://wikimediafoundation.org/2018/08/20/how-many-wikipedia-references-are-available-to-read/>.

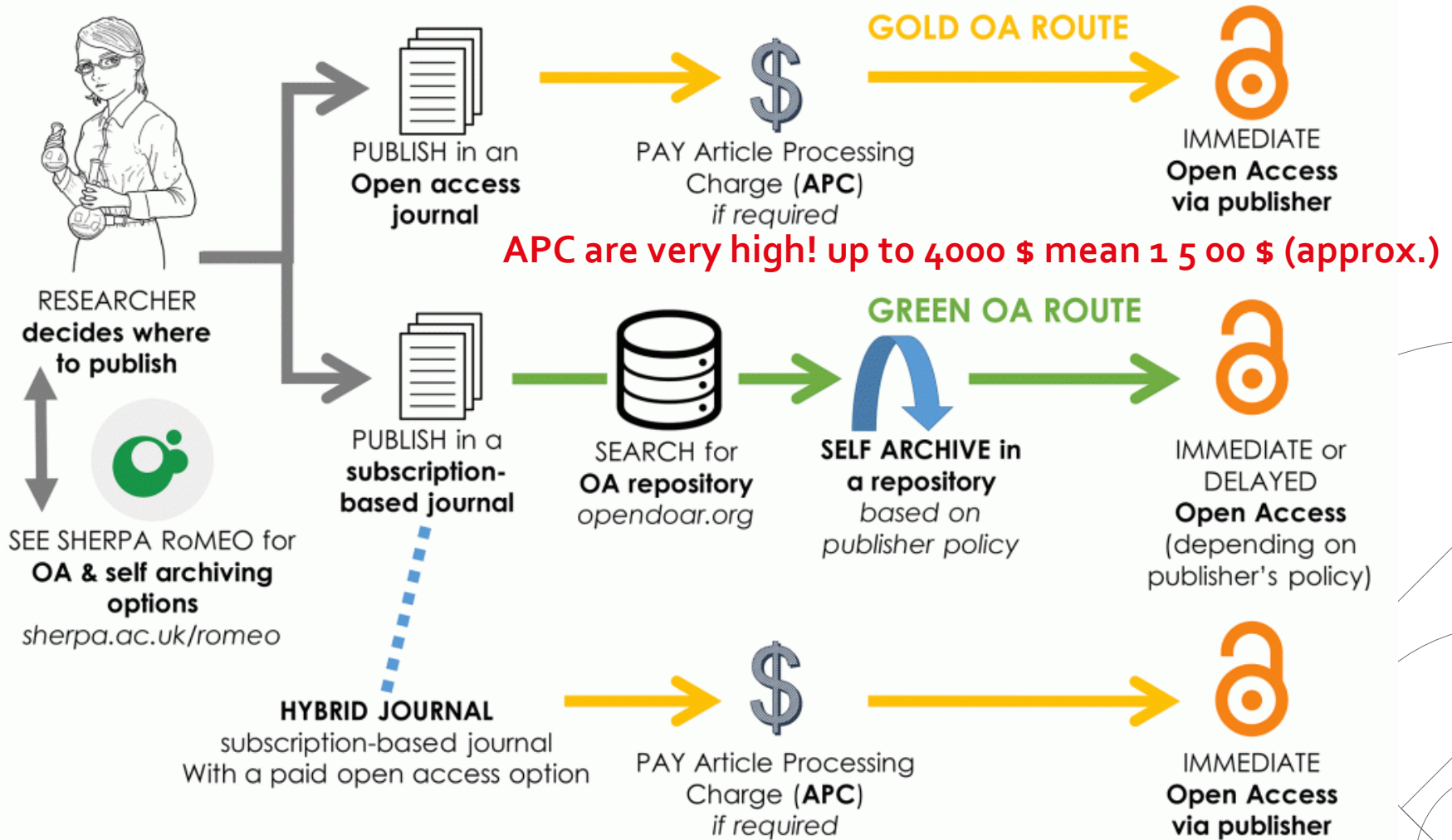
Open Access (OA)

Refers to research outputs which are distributed online and free of cost or other barriers, and possibly with the addition of a **Creative Commons license** to promote reuse. [Wikipedia]

All forms of published research output : peer-reviewed and non peer-reviewed academic journal articles, conference papers, theses, book chapters, and monographs.



Open Access Publishing



Digital Repository Fiocruz Brasil

An institutional repository

Green OA
Route



Fale com a Fiocruz

FUNDAÇÃO OSWALDO CRUZ



Repositório Institucional da Fiocruz

↑ Página inicial

Navegar ▾

Ajuda

Idioma ▾

Entrar: ▾



Autoarquivamento

Área Pessoal

Comunidades & Coleções

Pesquisar no ARCA



Novos documentos depositados

RSS 1.0

RSS 2.0

RSS

Louis Figuiet en Brasil: Algunas consideraciones acerca de los vulgarizadores de las ciencias en el último cuarto del siglo XIX

ARCA

- ▶ Sobre
- ▶ Equipe
- ▶ Termos de uso
- ▶ FAQ
- ▶ Fanpage Arca

Acesso Aberto

- ▶ O que é
- ▶ Glossário
- ▶ Direitos Autorais
- ▶ Política

Visualização de Dados



▶ Veja os dados do Arca de forma dinâmica

1284 revistas científicas 47 282 fascículos 611 925 artículos a texto completo

El portal de revistas de y para ciencias sociales y humanidades

Artículos



La ciencia que no se ve no existe



Colecciones de revistas



Autores Redalyc



Editores Redalyc



Sistema de Marcación XML-JATS Marcalyc

Gold OA Route



Redalyc comparte los principios de DORA y la Declaración de México

Sobre o SciELO

- Sobre o SciELO
- Indicadores Bibliométricos
- Acesso via OAI e RSS

Pesquisa artigos

método Entre com uma ou mais palavras onde
integrada Regional

Rede SciELO

coleções de Livros

SciELO Livros

coleções de Periódicos

- África do Sul
- Argentina
- Bolívia
- Brasil
- Chile
- Colômbia
- Costa Rica
- Cuba
- Espanha
- México
- Peru
- Portugal
- Uruguai
- Venezuela
- Saúde Pública

em desenvolvimento

- Equador
- Paraguai

Lista periódicos

Pesquisa periódicos

Entre com uma ou mais palavras

Por ordem alfabética - todos

A B C D E F G H I J K L M N O P
Q R S T U V W Y Z

Por assunto - todos

- Ciências Agrárias
- Ciências Biológicas
- Ciências da Saúde
- Ciências Exatas e da Terra
- Ciências Humanas
- Ciências Sociais Aplicadas
- Engenharias
- Linguística, Letras e Artes

Por publicador - todos

A B C D E F G H I J L M N O P R

SciELO em números

- Uso do site
- Citações
- Co-autoria

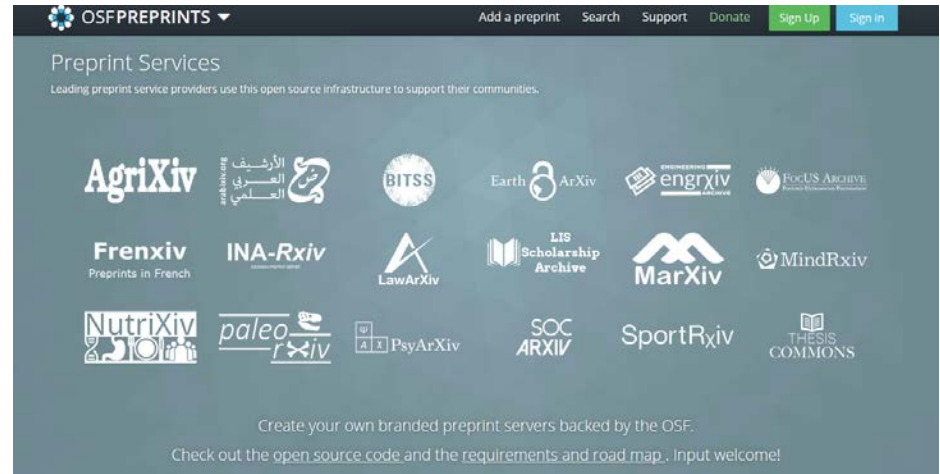
1285 Periódicos ativos
52356 Fascículos
745182 Artigos
16943454 Citações

Novos

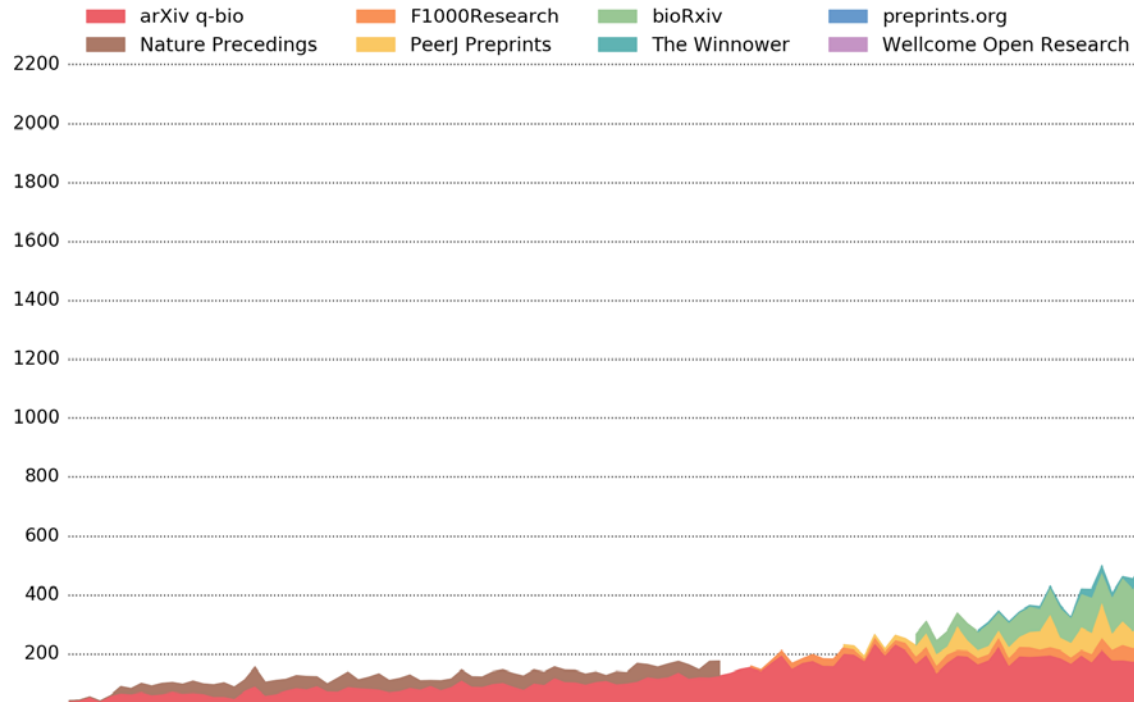
Última atualização - 04/set/2017

185 Fascículos
Por assunto
Por coleção

Preprints in development

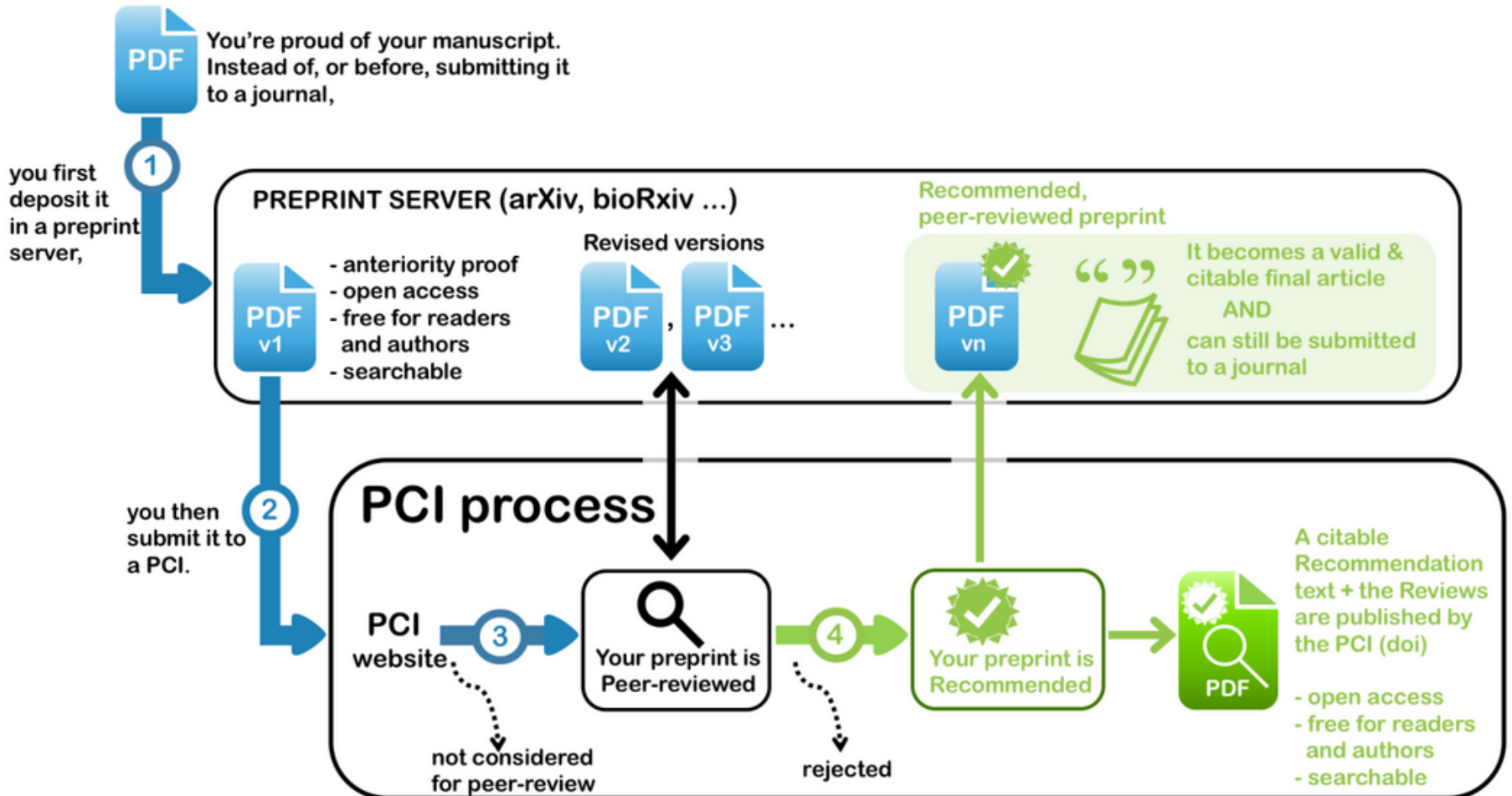


Preprints per Month



Peer Community In, Innovative platform

<https://peercommunityin.org/>



Research Data

Recorded factual material commonly accepted in the scientific community as necessary to validate research findings."

Research data covers a broad range of types of information and digital data can be structured and stored in a variety of file formats.

Properly managing data (and records) **does not necessarily equate to sharing or publishing that data.**

EXAMPLES OF RESEARCH DATA

Some examples of research data:

Documents (text, Word), spreadsheets

Questionnaires, transcripts, codebooks

Audiotapes, videotapes

Photographs, films

Protein or genetic sequences

...

Database contents (video, audio, text, images)

Models, algorithms, scripts

<https://www.lib.ncsu.edu/data-management/define>



Why share data ?

As Open as possible, As closed as necessary

- ◆ Reproducibility of results
- ◆ Enhanced citation : dataset can be cited as a product of research
- ◆ Respond to requirements of Funders or Publishers (Elsevier..)
- ◆ Reusability for others researchers (or anyone)
- ◆ Data Preservation

Data management requires forward planning



A **data management plan** describes how you will collect, organise, manage, store, secure, back- up, preserve, and where applicable, share your research data

<https://library.nuigalway.ie/digitalscholarship/researchdata/researchdatamanagement/>

Recent uploads

October 16, 2018 (v2.1) Dataset Open Access

SemFi - Finnish Semantic Database with Syntactic Relations


Hämäläinen, Mika

SemFi is a semantic database for Finnish in which the words are linked to each other by the syntactic relations and their frequency in a big corpus. SemFi is based on the syntactic bigrams of The Finnish Internet Parsebank provided by Turku University. The semfi.db file is an SQLite database and...

Uploaded on October 16, 2018
2 more version(s) exist for this record


View

Zenodo now supports usage statistics!



[Read more](#) about it, in our newest blog post.

Using GitHub?



Just [Log in](#) with your GitHub account and [click here](#) to start preserving your repositories.

Zenodo in a nutshell

- **Research. Shared.** – all research outputs from across all fields of research are

Dataverse

October 15, 2018 (v2.6.1) Software Open Access

sjPlot - Data Visualization for Statistics in Social Science.

Daniel Lüdecke

General Removed defunct functions. Deprecated sjt.lm(), sjt.glm(), sjt.lmer() and sjt.glmr() are now deprecated. Please use tab_model() instead. Changes to functions Arguments dot.size and line.size in p effects and diagnostic plots. plot_model() now uses a...

Uploaded on October 15, 2018
3 more version(s) exist for this record

View

October 1, 2018 (v2.0) Preprint Open Access

33 Installations

2,872 Dataverses

52,449 Datasets

4,102,900 Downloads



Depositing data in a repository makes it easier to preserve, discover and reuse data

◆ Repositories are scanned by specific search tools



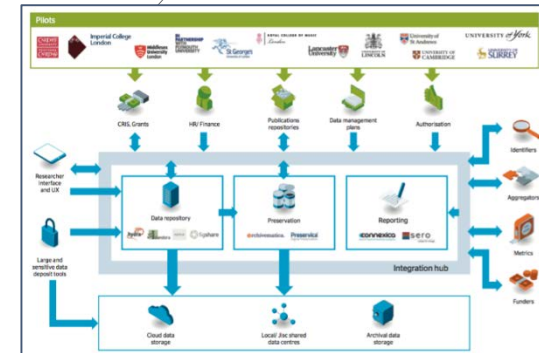
- ◆ Datacite search
- ◆ Data Citation Index (Thomson Reuters)
- ◆ Data search (Elsevier)
- ◆ ...



DataSearch Beta

◆ Can disseminate their data via the OAI-PMH exchange protocol

◆ and harvested by catalogs, integrators, European data infrastructures ... more and more numerous



Research Data Shared Service (JISC)

D'après Cocaud et Aventurier 2017

Filter

- Subjects ⊕
- Content Types ⊕
- Countries ⊕
- AID systems ⊕
- API ⊕
- Data access ⊕
- Data access restrictions ⊕
- Database access ⊕
- Database licenses ⊕
- Data licenses ⊕
- Data upload ⊕
- Data upload restrictions ⊕
- Enhanced publication ⊕
- Institution responsibility type ⊕
- Institution type ⊕
- Keywords ⊕
- Metadata standards ⊕
- PID systems ⊕
- Provider types ⊕
- Quality management ⊕
- Repository languages ⊕
- Software ⊕
- Syndications ⊕
- Repository types ⊕
- Versioning ⊕

reef

← Previous 1 Next →

Found 6 result(s)

CoRIS

[Coral Reef Information System](#)

Subject(s)

Biology Atmospheric Science and Oceanography Life Sciences Geosciences (including Geography) Natural Sciences

Content type(s)

Standard office documents Images Scientific and statistical data formats Raw data

Country

United States

NOAA's Coral Reef Information System (CoRIS) is a web-based information portal that provides access to NOAA coral reef information and data products with emphasis on the U.S. states, territories and remote island areas. NOAA Coral Reef activities include coral reef mapping, monitoring and assessment; natural and socioeconomic research and modeling; outreach and education; and management and stewardship.

Temperate Reef Base

[Temperate Reef Base Portal](#)

Subject(s)

Plant Ecology and Ecosystem Analysis Geosciences (including Geography) Natural Sciences Plant Sciences Biology Life Sciences

Atmospheric Science and Oceanography Oceanography

Content type(s)

Raw data Images Scientific and statistical data formats Plain text Databases

Country

Australia International

Additional services
Open/restricted/closed
Identifier (DOI)
Data policy
Certification



Datacite : tools and Schema



Find what you're looking for by searching millions of records with extensive, reliable metadata.



Share your data and reuse the data of others to create the highest impact in the research community.

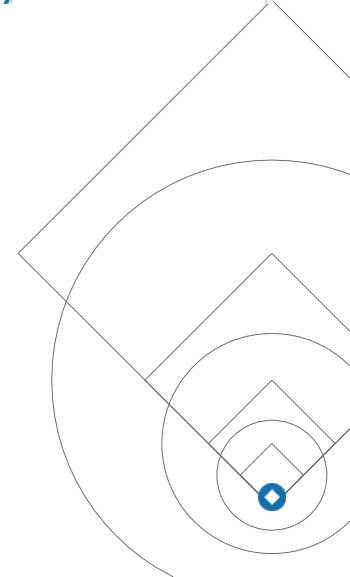
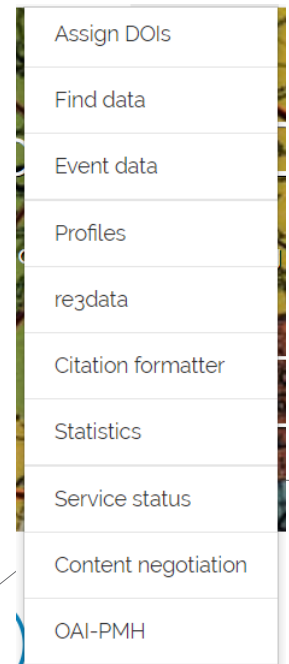


Cite your research sources with confidence, and receive proper credit when your work is reused.



Connect your research – publications, datasets, software, authors, institutions, and funding data all in one place.

Metadata schema for datasets
Provide DOI for data set
Citation format
and much more



Fair Principles

- Describe your data in a data repository
- Apply a persistent identifiers

Findable

- Consider what will be shared
- Obtain participant consent & perform risk management

Accessible

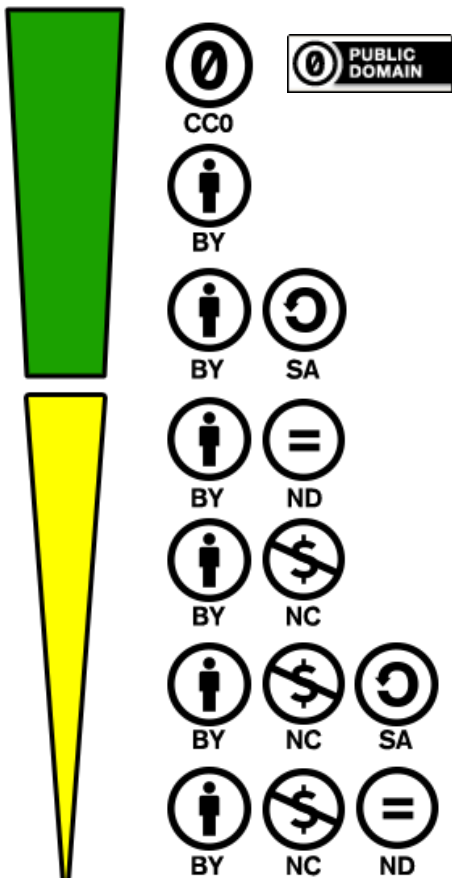
Interoperable

- Use open formats
- Consistent vocabulary
- Common metadata standards

Reusable

- Consider permitted use
- Apply appropriate licence

MOST OPEN




LEAST OPEN

dd/mm/yyyy

Share your work

Use Creative Commons to help share your work.

CC copyright licenses provide [a simple, standardized way to give your permission to share and use your creative work](#)— on conditions of your choice.

-  [Attribution \(by\)](#)
-  [Share Alike \(sa\)](#)
-  [NonCommercial \(nc\)](#)
-  [NoDerivatives \(nd\).](#)

Open Science Data Policy

An example from the FWF Austrian Science Fund

- ◆ All research should be findable accessible interoperable and reusable fulfill the FAIR principles and the following criteria need to apply
- ◆ The data repository must be certified. While certification in **CoreTrustSeal** is explicitly recommended, certification in **re3data** is required
- ◆ Data should be deposited in such a way that can be re-used without restrictions (eg CC0, CC BY, etc.)
- ◆ Deposited datasets have to be citable by means of a persistent identifier (.e. g; DOI)

source https://twitter.com/tonyR_H/status/1057212903118520320





DORA

Dora Declaration on Research Assessment (since 2012)

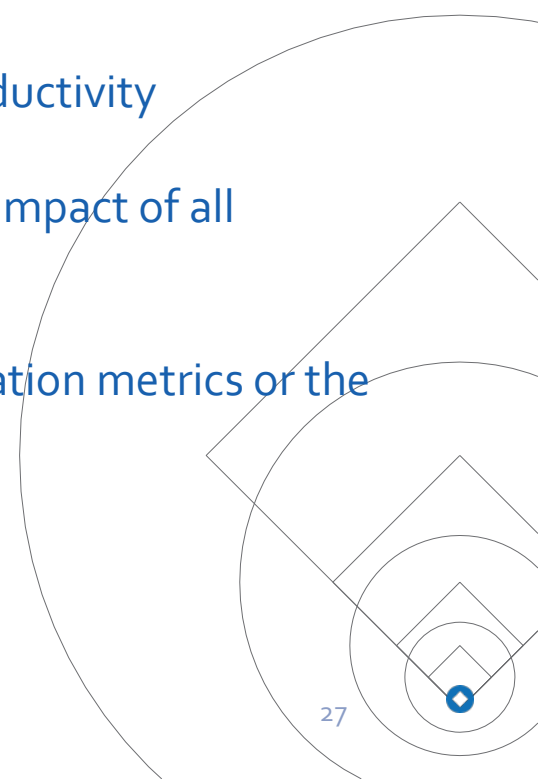
Signed by 705 organisations and more than 13 000 individuals

Do not use journal-based metrics, such as Journal Impact Factors for article or individual evaluation

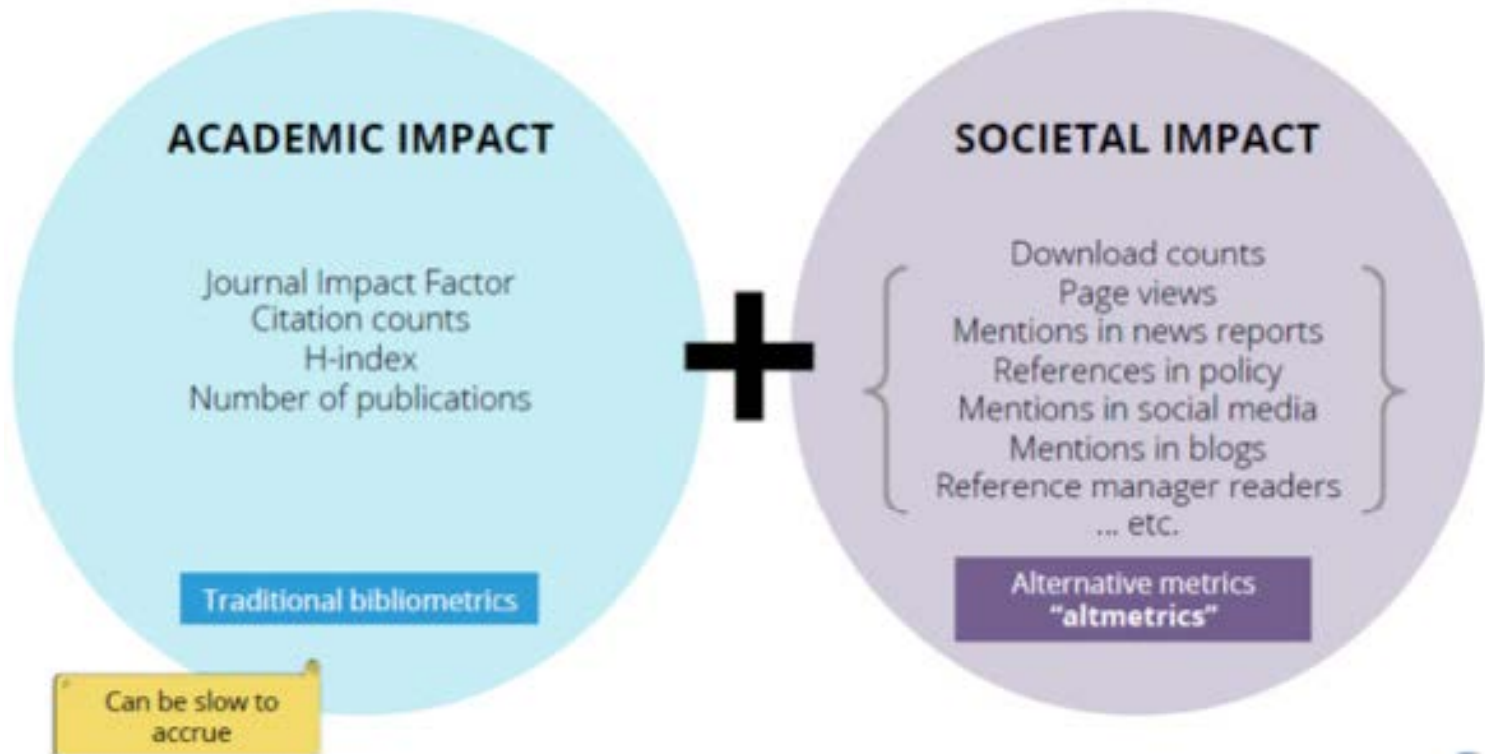
Be explicit about the criteria used in evaluating the scientific productivity

For the purposes of research assessment, consider the value and impact of all research outputs (including datasets and software)

Scientific content of a paper is much more important than publication metrics or the identity of the journal in which it was published.



Altmetrics : alternatives metrics



<http://www.ucd.ie/research/portal/researchperformanceself-servicereporting/altmetrics/>

Conclusion

Open Science is not about Dogma. It is about greater efficiency and productivity more transparency and better response to interdisciplinary needs...

Open science involves a lot of technical components and systems

But key factors are human factors as collaboration between services (Libraries, law, ICT...

References

- ◉ Cocard, Sylvie, and Pascal Aventurier. 2017. "Les Entrepôts de Données." presented at the ANF Renatis, Nancy.
- ◉ Desconnets, Jean-Christophe, Pascal Aventurier, Caroline Doucouré, and Stéphane Banon. 14:04:45 UTC. "Entrepôt de Données Ouvertes IRD. Un Service En Ligne Pour l'ouverture..." Education. https://www.slideshare.net/IST_IRD/entrepot-de-donnees-ouvertes-ird-un-service-en-ligne-pour-louverture-et-le-partage-des-donnees-scientifiques-au-sud-2018.
- ◉ Jones, Sarah. 2011. "How to Develop a Data Management and Sharing Plan," 8. <http://www.dcc.ac.uk/resources/how-guides/develop-data-plan>
- ◉ Rentier, Bernard. Bernard Rentier on Open Access: From Where? How? For What? And Where To? | Eurodoc." 2018. Accessed October 29, 2018. <http://eurodoc.net/news/2018/bernard-rentier-on-open-access-from-where-how-for-what-and-where-to>.
- ◉ "Data Management - H2020 Online Manual." 2018. http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm. "
- ◉ "Open Science and Its Role in Universities: A Roadmap for Cultural Change." 2018. LERU. Accessed October 29, 2018. <https://www.leru.org/publications/open-science-and-its-role-in-universities-a-roadmap-for-cultural-change>.
- ◉ *Open Science by Design: Realizing a Vision for 21st Century Research* n.d. THE NATIONAL ACADEMIES PRESS.
- ◉ "Open Science | What Is Open Science?" n.d. Accessed October 31, 2018. <https://www.openscience.nl/en/open-science/what-is-open-science/index>.
- ◉ "LEARN Project – LEARN Project: Leaders Activating Research Networks." n.d. Accessed October 31, 2018. <http://learn-rdm.eu/en/about/>.