



Mendeley Data

Research data sharing:
stick or carrot?

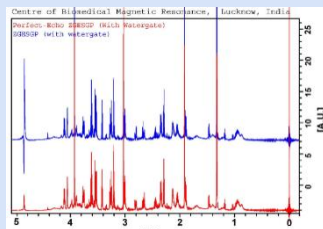
Wouter Haak – VP Research Data Management

w.haak@Elsevier.com

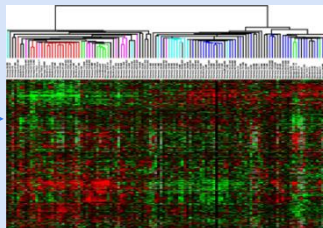


When talking about research data, we mean:

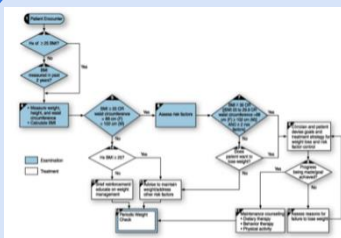
All forms of data



Raw data



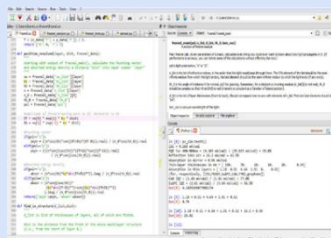
Processed data



Protocols, methods, workflows



Machine & environment settings



Scripts, analyses & algorithms

FAIR Data Principles



Findable



Accessible



Interoperable



Reusable

← → ↺ 🏠

data-science.nih.gov/covid-19-open-access-resources

☆ 📷 ⌂ 🔍 👤

📱 Apps 📁 RDM demo 📁 RDM Sales 📁 RDM general 📁 Elsevier general 🌐 📄 RM Sales Team - Kn... ☁️ RDM - OneDrive W... 📄 CPC Master sheet...

🇺🇸 U.S. Department of Health & Human Services 🏛️ National Institutes of Health 📄 Division of Program Coordination, Planning, and Strategic Initiatives (DPCPSI)

NIH

National Institutes of Health

Office of Data Science Strategy

f

🐦

Home

Strategic Plan

Resources

Research Funding

About

COVID-19 is an emerging, rapidly evolving situation.

Get the latest public health information from CDC: <https://www.coronavirus.gov>
Get the latest research information from NIH: <https://www.nih.gov/coronavirus>
NIH and other federal agencies have made COVID-19 data available through several [Open-Access Data and Computational Resources](#)

Open-Access Data and Computational Resources to Address COVID-19



Open-Access Data and Computational Resources to Address COVID-19

COVID-19 open-access data and computational resources are being provided by federal agencies, including NIH, public consortia, and private entities. These resources are freely available to researchers, and this page will be updated as more information becomes available.

The Office of Data Science Strategy seeks to provide the research community with links to open-access data, computational, and supporting resources. These resources are being aggregated and posted for scientific and public health interests. Inclusion of a resource on this list does not mean it has been evaluated or endorsed by NIH.

To suggest a new resource, please send an email with the name of the resource, the website, and a short description to datascience@nih.gov.

[See Computational Resources](#)
[See Supporting Resources](#)

Search:

Filter by Data Type

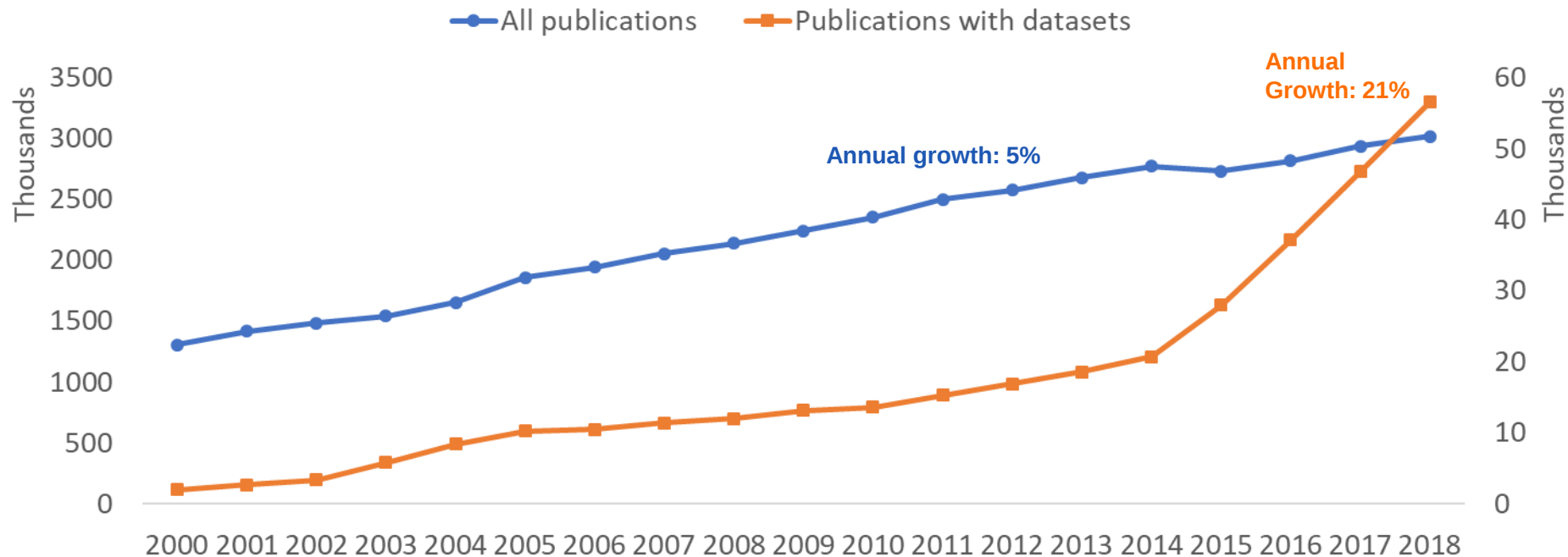
- Any -

Apply

Resource	Resource Description	Data Type	NIH Funded
----------	----------------------	-----------	------------

In times of Covid19 – data sharing is even more important

RDM adoption is growing very fast worldwide



19.06.2020

Source: Mendeley Data Monitor analysis of Scopus, Scholix, SciVal, 5 year data 2014-2018
extracted on August, 2019 – Annual Growth = Compound Annual Growth Rate (CAGR)

Stick or carrot?

Three trends are coming together, leading to an acceleration of RDM practices

1) Computational:

- Faster, easier, cheaper, more computational methods of doing science
- Coming of age of analytics yield new layers of insight on same data

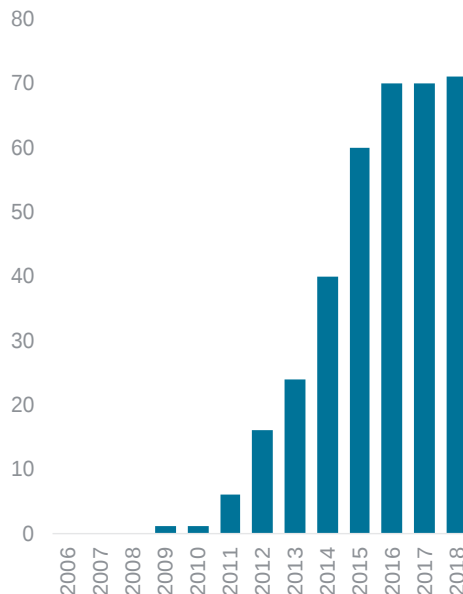
2) Funding bodies & institutions:

- Funding agencies driving data sharing mandates aiming to improve re-use of data and reproducibility of research
- Which is followed by institutions adopting data management policies and plans

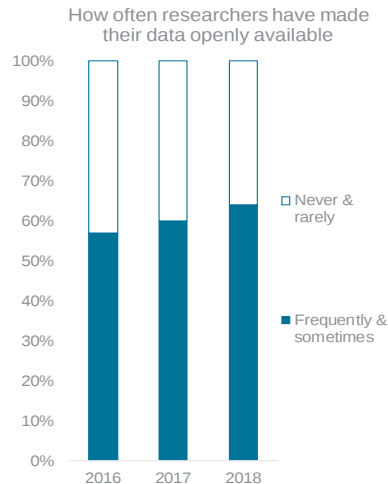
3) Researchers:

- a new generation of researchers, more focused on data and data sharing
- “Research Data is a first class citizen”

Example of trend 2
Number of UK institutions
adopting a data management
policy



Examples of trend 3:



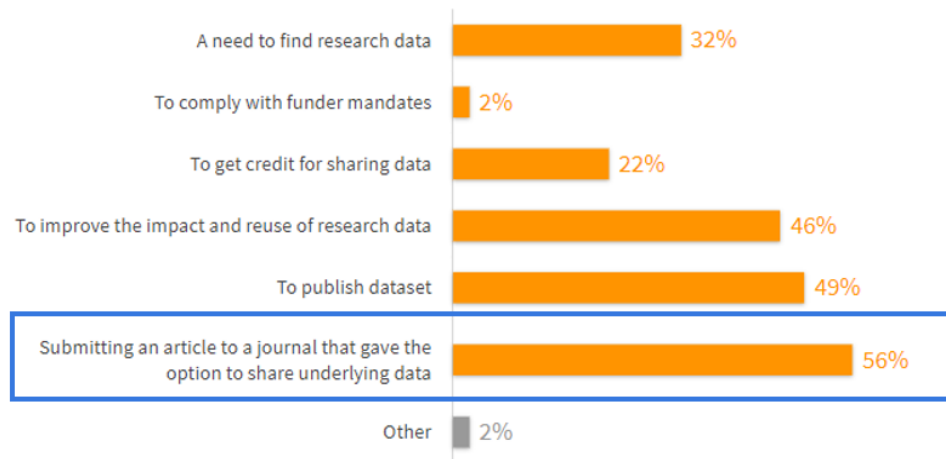
How researchers value data citations

Date of first peer reviewed publication	Respondents who value a data citation same or more than an article citation	Respondents who value a data citation less than an article citation
<1990	64%	36%
90s	65%	35%
00s	60%	40%
10s	72%	28%



An example: why do researchers use the Elsevier data sharing environment?

“What prompted you to use Mendeley Data?”

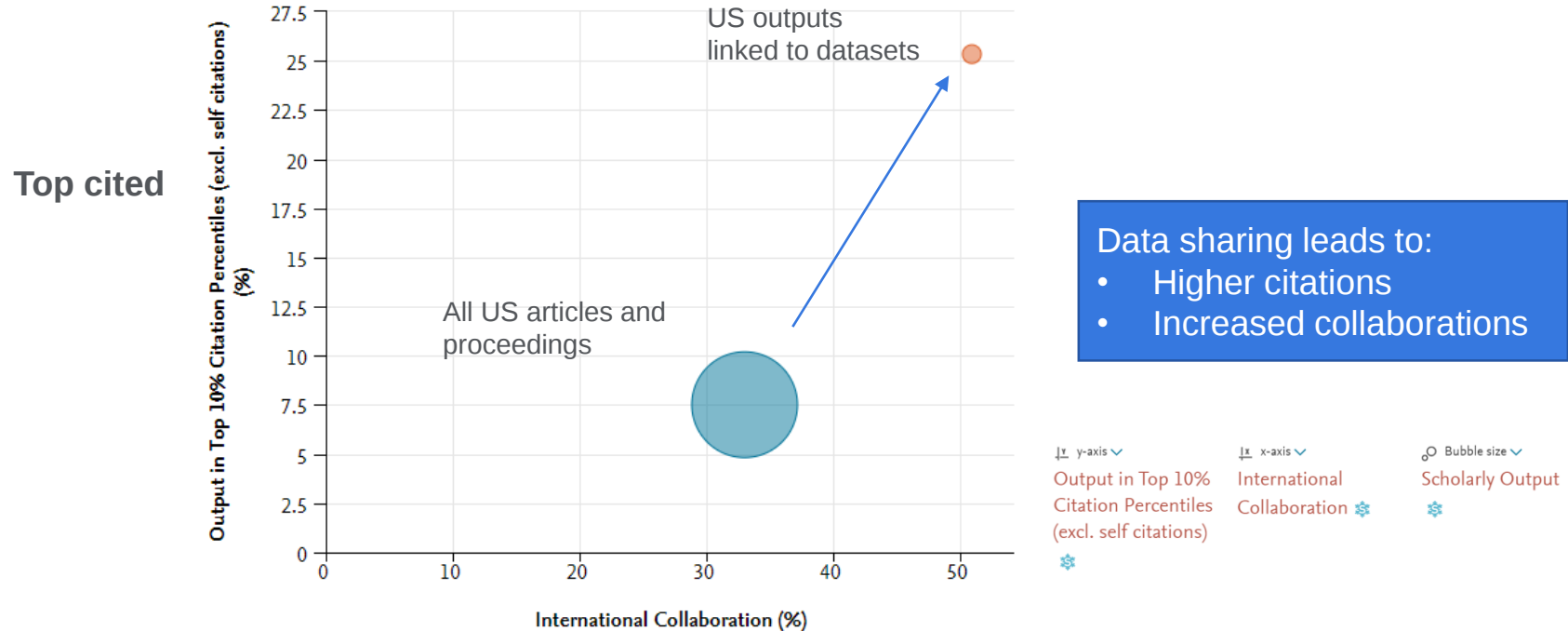


Source: TechValidate survey of 387 users of Mendeley Data

19.06.2020

“So what?”

The impact of sharing data in (for example) the US

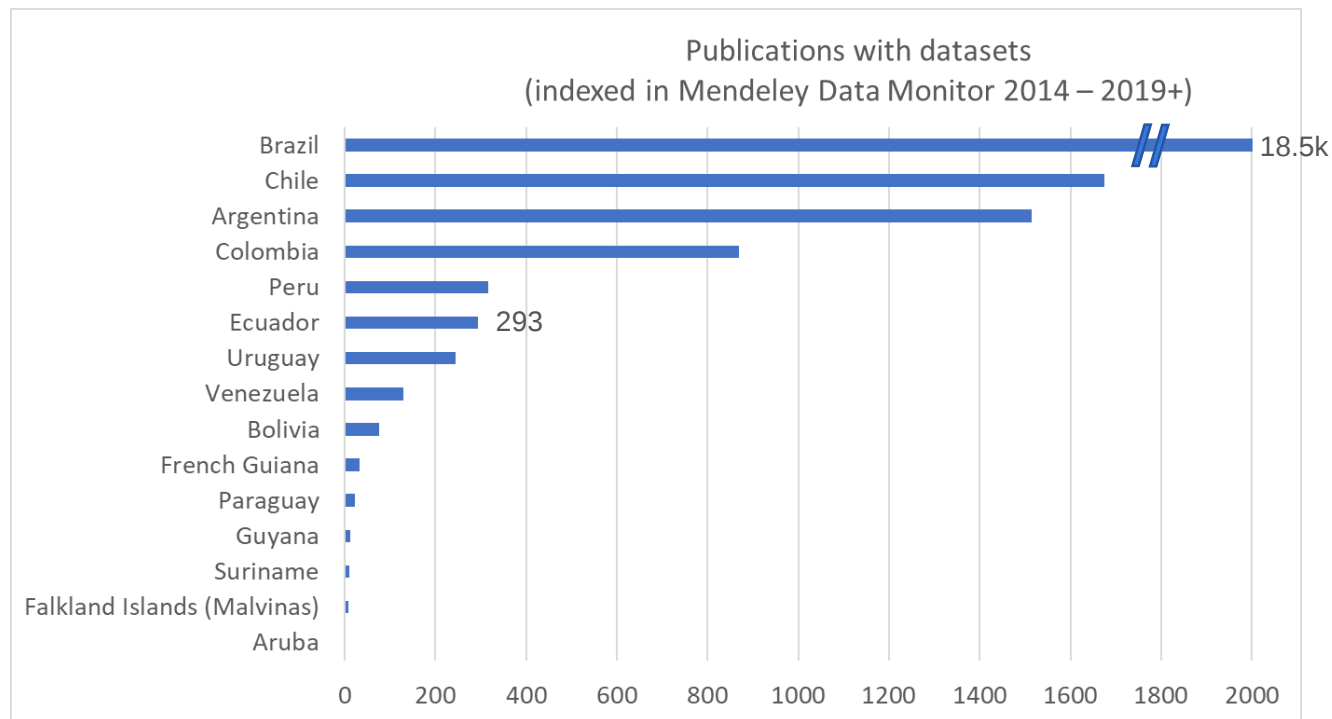


Scival: research area: "Publications with datasets indexed in Mendeley Data Monitor (2014 – 2019+)"

Analysis in Trends module, selecting the specific country and institution; exported these articles as a separate set and benchmarked against overall inst output in Benchmarking module

Research area is based on a cross section between Scopus (affiliation), Scholix (article-data links), as stored in Mendeley Data Monitor database

Data sharing in Latin America is happening!

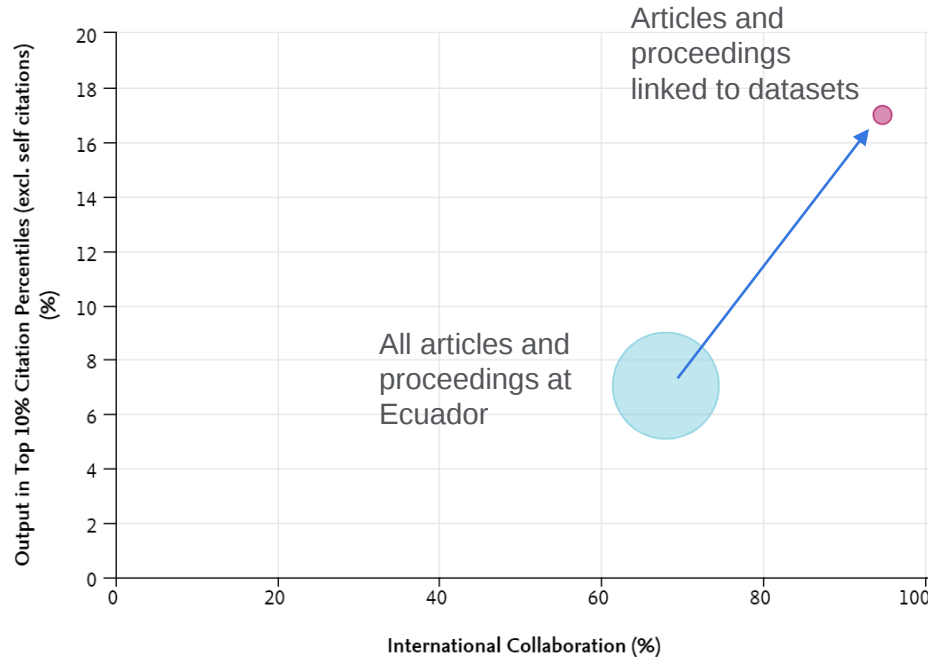


Scival: research area: "Publications with datasets indexed in Mendeley Data Monitor (2014 – 2019+)"

Analysis in Trends module, selecting the specific country and institution; exported these articles as a separate set and benchmarked against overall inst output in Benchmarking module

Research area is based on a cross section between Scopus (affiliation), Scholix (article-data links), as stored in Mendeley Data Monitor database

Impact of sharing research data: Ecuador



Data sharing leads to:

- Higher citations
- Increased collaborations

y-axis

Output in Top 10%
Citation Percentiles
(excl. self citations)



x-axis

International
Collaboration



Bubble size

Scholarly Output

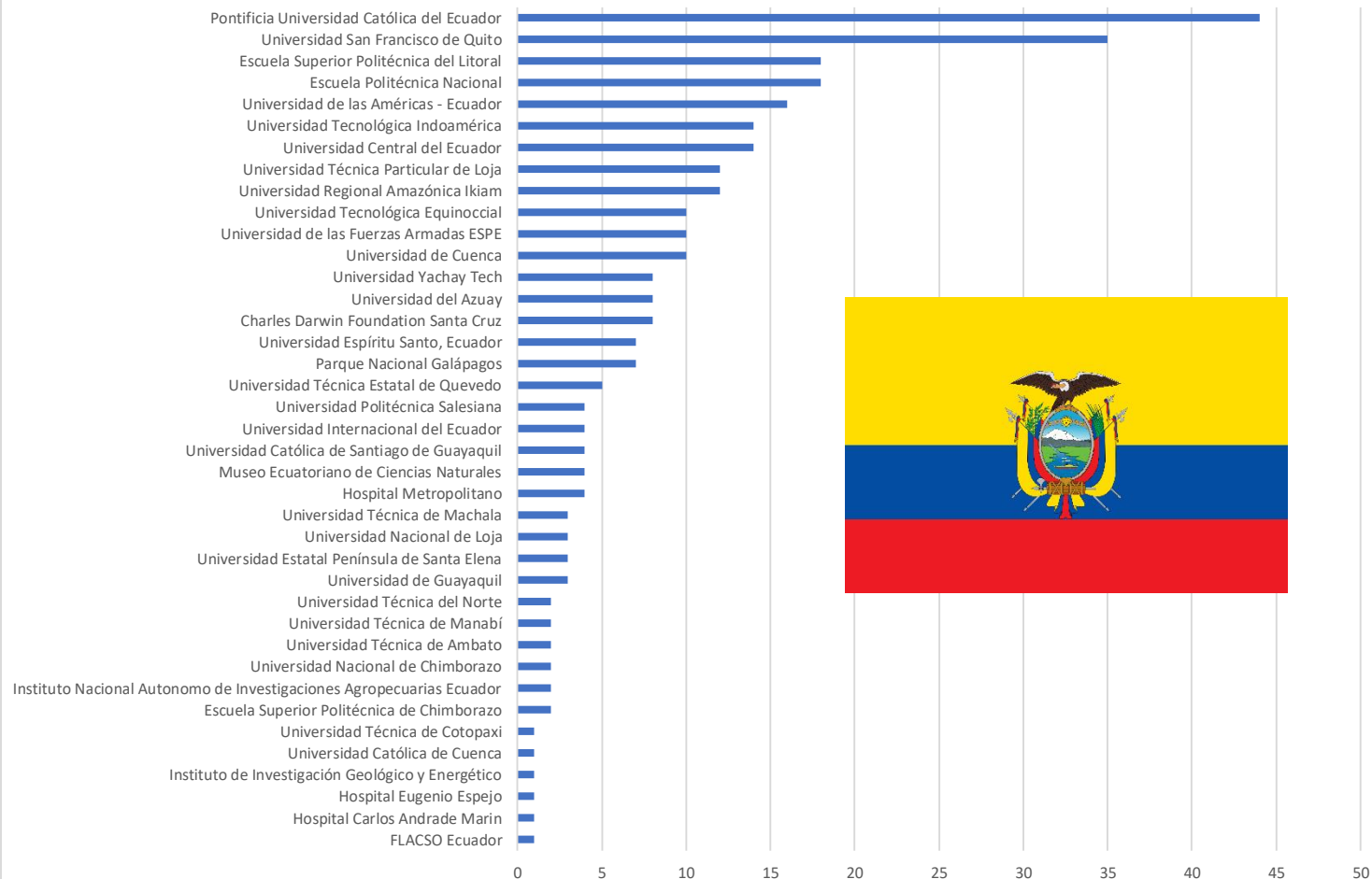


Scival: research area: "Publications with datasets indexed in Mendeley Data Monitor (2014 – 2019+)"

Analysis in Trends module, selecting the specific country and institution; exported these articles as a separate set and benchmarked against overall inst output in Benchmarking module

Research area is based on a cross section between Scopus (affiliation), Scholix (article-data links), as stored in Mendeley Data Monitor database

Publications with datasets (indexed in Mendeley Data Monitor 2014 – 2019+)



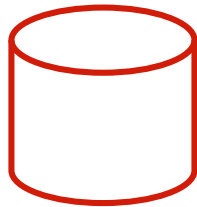
“How about my institution?”

Assessing RDM for an institution

Research data location	Purpose	Considerations
Data shared on generalist or domain repositories 	• A secure place for all researchers to store and share data • Primarily drives research re-usability and reproducibility • >90% of public research data	• Do you have a data policy or do you require data management plans? Is it successful? • Do you track, report, and get credit for this data? • Can your researchers find this data? How?
Data shared on your institutional (data) repository 	• A secure place for your researchers to store and share data • Primarily drives institutional recognition, as well as research re-usability and reproducibility • <10% of public research data	• Does your (general-purpose) institutional repository meet the needs for data sharing? • Do you know the benefits of a data-specific institutional repository? • Do you showcase all data outputs?
Data shared privately by or with your researchers 	• 'Active' data in research lifecycle • Collaboration on sensitive and local data • Help researchers make data FAIR	• Do you have visibility of active data? • How do you help researchers manage and collaborate with others on active data? • Do you have Data Stewards or Champions?

Organize, curate and report on public research data to enhance visibility and compliance

It's difficult for institutions to have full visibility of thousands of open data repositories and monitor which ones their researchers are using.



~90%

of public research data
is hosted on external
generalist or domain
specific repositories



~10%

of public research
data is hosted on
institutional
repositories

How to add a bit more 'carrot' in order to share research data



RESEARCH DATA MANAGEMENT LIBRARIAN ACADEMY [RDMLA]

Started Oct 7, 2019

COURSE DATE:

Open

DURATION:

Ongoing

COMMITMENT:

1.5-2 hrs/unit

REQUIREMENT:

None

COURSE TYPE:

Self-paced

CREDENTIAL:

Certificate; Cont Ed
Credits (CEUs)

You are enrolled.

Elsevier's Mendeley Data: plug-and-play solution to share data

Mendeley Data Monitor & Search

A tool to track datasets published by researchers both within and outside the institution, to facilitate compliance with funders' mandates and enable reporting and showcasing.

As well as a comprehensive data search engine including 20+ million datasets indexed from 1000s of data repositories.

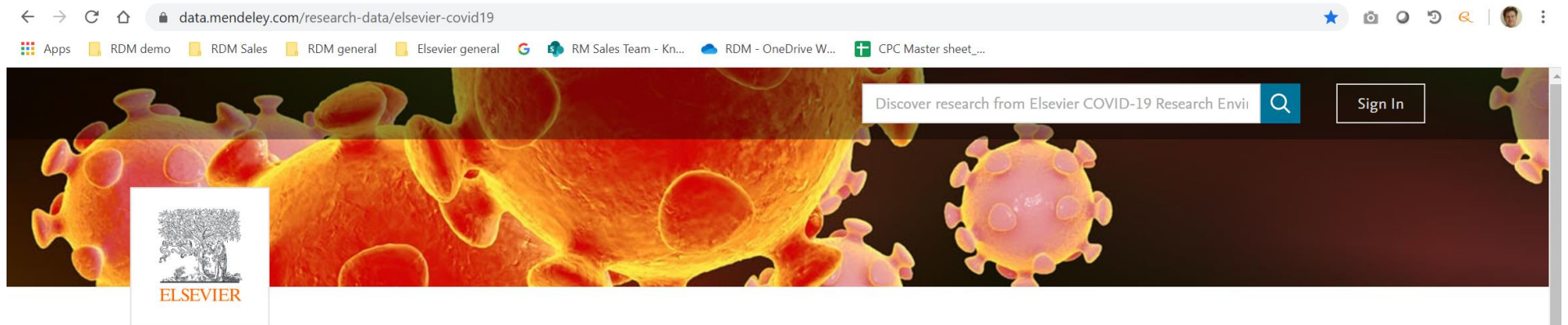
Mendeley Data Repository & Manager

A repository specialized for research data, to store and share datasets following the FAIR data principles.

A collaborative team workspace where researchers can share project data, integrate external data sources securely and prepare datasets for publication.



The Mendeley Data platform gives you full control to successfully drive forward your institutional Research Data Management (RDM) policy



Elsevier COVID-19 Research Environment



[Advanced search](#)

Filter Results

DATA TYPES

☒ Collection (4)

SOURCE

☒ Mendeley Data (4)

4 results

 **COVID-19: Public health, and societal and psychological impacts datasets**

Contributors: Elsevier Team

Date: 2020-04-13

We selected Public health, and societal and psychological impacts datasets indexed by the Mendeley Data Search engine on the 2019-present COVID-19 / Coronavirus pandemic. The aim was to make it easier to find potentially relevant datasets for this specific topic

Data Types:

Collection

 **COVID-19: Epidemiology & infectious modelling datasets**



Sign in

Create account

Download

Reference Management Research Network **Datasets** Careers Funding

Find Research Data My Datasets New Dataset FAQ

Protein regulatory relationships in COVID19

Published: 25 Apr 2020 | **Version 2** | DOI: 10.17632/3pm7yy4xj9.2

Contributor(s): [Finlay Maclean](#), [Helena Deus](#), [Tony Scerri](#)

Description of this data

This dataset was extracted from the Elsevier Pathway Studio, a tool that helps scientists analyze experimental data to answer biologically meaningful questions. The dataset itself consists of biological relationships between diseases (MERS and SARS), proteins and molecules. The relationships are of various types including Regulation, Target, Molecular Transport, etc. You can find a mapping of the relationship name to a description on this support page: https://service.elsevier.com/app/answers/detail/a_id/3014/supporthub/pathway/

The source of these relationships are life sciences and biomedical articles, from various publishers. We make use of taxonomies, curated and maintained by subject matter experts, to extract the right terms from text and map them to the correct identifiers. Subject matter experts have also helped us create the rules and information extraction patterns to optimize the extraction of relationships from text.

At last, the pubmed identifiers from which the relationships were extracted are also part of the dataset. The .cypher and .json files can be imported into the graph database neo4j. The .csv files can be used to import into other systems.

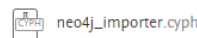
Experiment data files

Download all files (4)



edges.csv

19 MB Cite



neo4j_importer.cypher

45 MB Cite

Latest version

Version 2

2020-04-25

Published: 2020-04-25

DOI: 10.17632/3pm7yy4xj9.2

Cite this dataset

Maclean, Finlay; Deus, Helena; Scerri, Tony (2020), "Protein regulatory relationships in COVID19 ", Mendeley Data, v2

<http://dx.doi.org/10.17632/3pm7yy4xj9.2>

Statistics

Views: **791**

Downloads: **73**

Previous versions

Version 1

2020-04-24

Mendeley Data Search

A powerful search engine specialized for data

Mendeley Data Search indexes 21+ million datasets from 1,700+ domain specific, generalist and institutional repositories

Data search accuracy through deep indexing and advanced search

Mendeley Data Search has the power to index both the metadata and actual data within a dataset. This, combined with advanced query processing and search capabilities, ensures that researchers get more accurate and comprehensive search results.





Find Research Data My Datasets New Dataset FAQ

("COVID-19" OR "2019-nCoV" OR "COVID19" OR "nCoV-19" OR "Sars-CoV-2")



[Advanced search](#)

Filter Results

7233 results

DATA TYPES

- ☐ Document (2923)
- ☐ Dataset (2218)
- ☐ Other (863)
- ☐ Video (470)
- ☐ Image (382)
- ☐ Software/Code (285)
- ☐ Collection (203)
- ☐ Tabular Data (116)
- ☐ File Set (37)
- ☐ Text (14)
- ☐ Interactive Resource (11)
- ☐ Sequencing Data (4)
- ☐ Event (1)
- ☐ Geospatial Data (1)
- ☐ Model (1)

REPOSITORY TYPES

- ☒ Data Repositories (7233)

SOURCE

- ☐ figshare Academic Research System (2360)
- ☐ Zenodo (1762)
- ☐ IEEE DataPort (366)
- ☐ Open Science Framework (310)
- ☐ Kaggle (296)
- ☐ Crossref Citations (294)
- ☐ Data Planet (282)
- ☐ Mendeley Data (210)
- ☐ Harvard Dataverse (105)



Coronavirus COVID-19 (2019-nCoV) Data Repository for Africa

Contributors: Marivate, Vukosi, Nsoesie, Elaine, Esube Bekele, Africa open COVID -19 data working group
Date: 2020-03-30

covid19... url:https://zenodo.org/communities/covid-19

Data Types:

Dataset

Export: [APA](#) [BibTeX](#) [DataCite](#) [RIS](#)



Coronavirus COVID-19 (2019-nCoV) Data Repository for Africa

Contributors: Marivate, Vukosi, Nsoesie, Elaine, Esube Bekele, Africa open COVID -19 data working group
Date: 2020-03-30

covid19... url:https://zenodo.org/communities/covid-19

Data Types:

Dataset

Export: [APA](#) [BibTeX](#) [DataCite](#) [RIS](#)



Medical Countermeasures Analysis of 2019-nCoV / SARS-CoV-2 for COVID-19

Contributors: Ricke, Darrell
Date: 2020-01-01

SARS-CoV-2 viral protein multiple sequence alignments and Flux variation analysis

Data Types:

Dataset

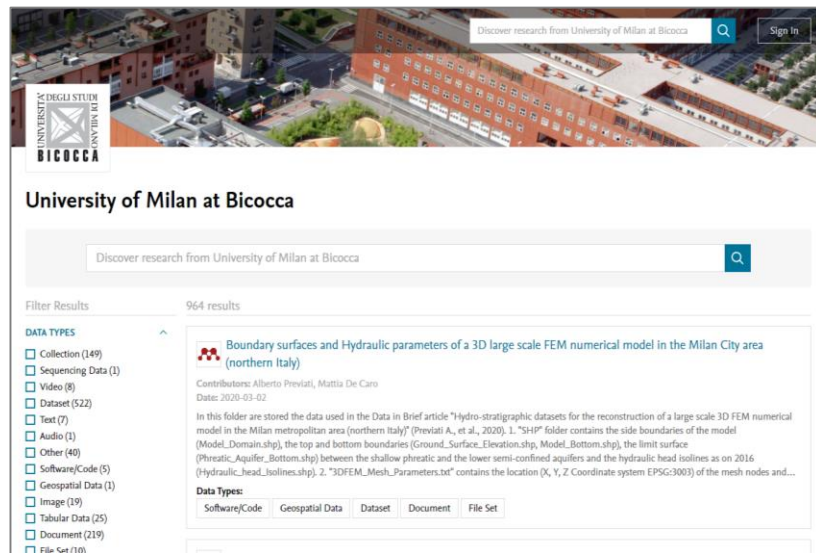
Export: [APA](#) [BibTeX](#) [DataCite](#) [RIS](#)

Showcase datasets to improve discoverability and reuse

for datasets in your institutional repository as well as for datasets published in one of the many external repositories

As an institution you can:

- **Get higher scores for grant applications** by promoting public research data output of your Institution
- **Improve your institutional research data visibility** by making it easy to find and reuse

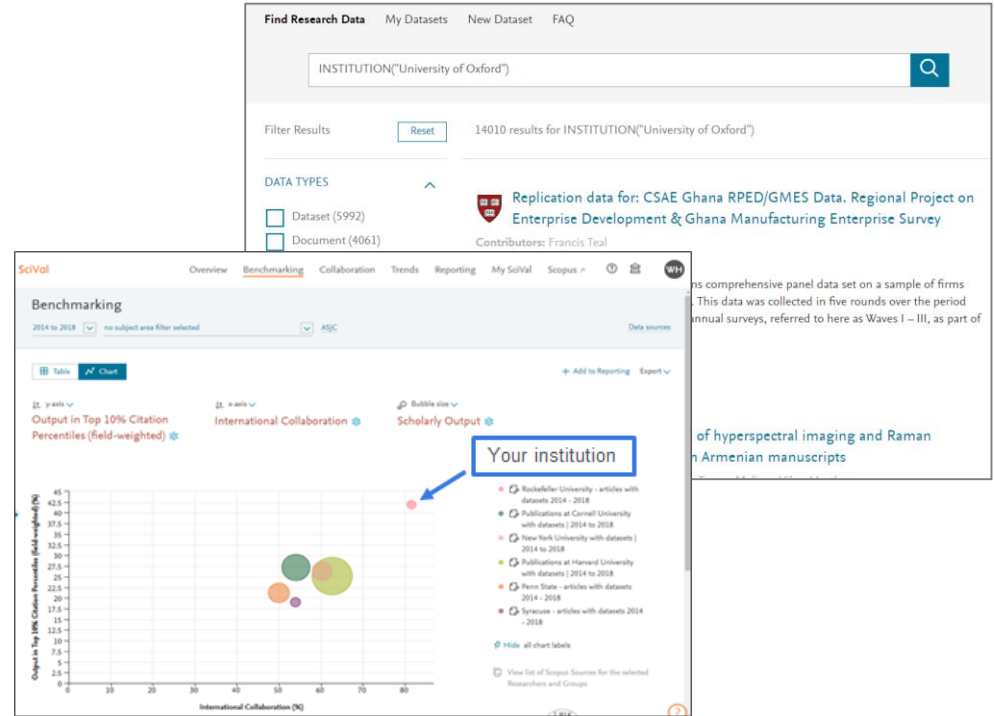


Mendeley Data Monitor

Compare and benchmark your research data output

As an institution you can **compare and benchmark your research data output** with other institutions

- via advanced data search queries, or
- by using Elsevier products featuring Mendeley Data Monitor integrations (e.g. SciVal)





Thank you

Wouter Haak – w.haak@Elsevier.com

