

Revelations about relations in connecting research: content types, data and identifiers



Rachael Lammey

Head of Community Outreach

Crossref

<https://orcid.org/0000-0001-5800-1434>



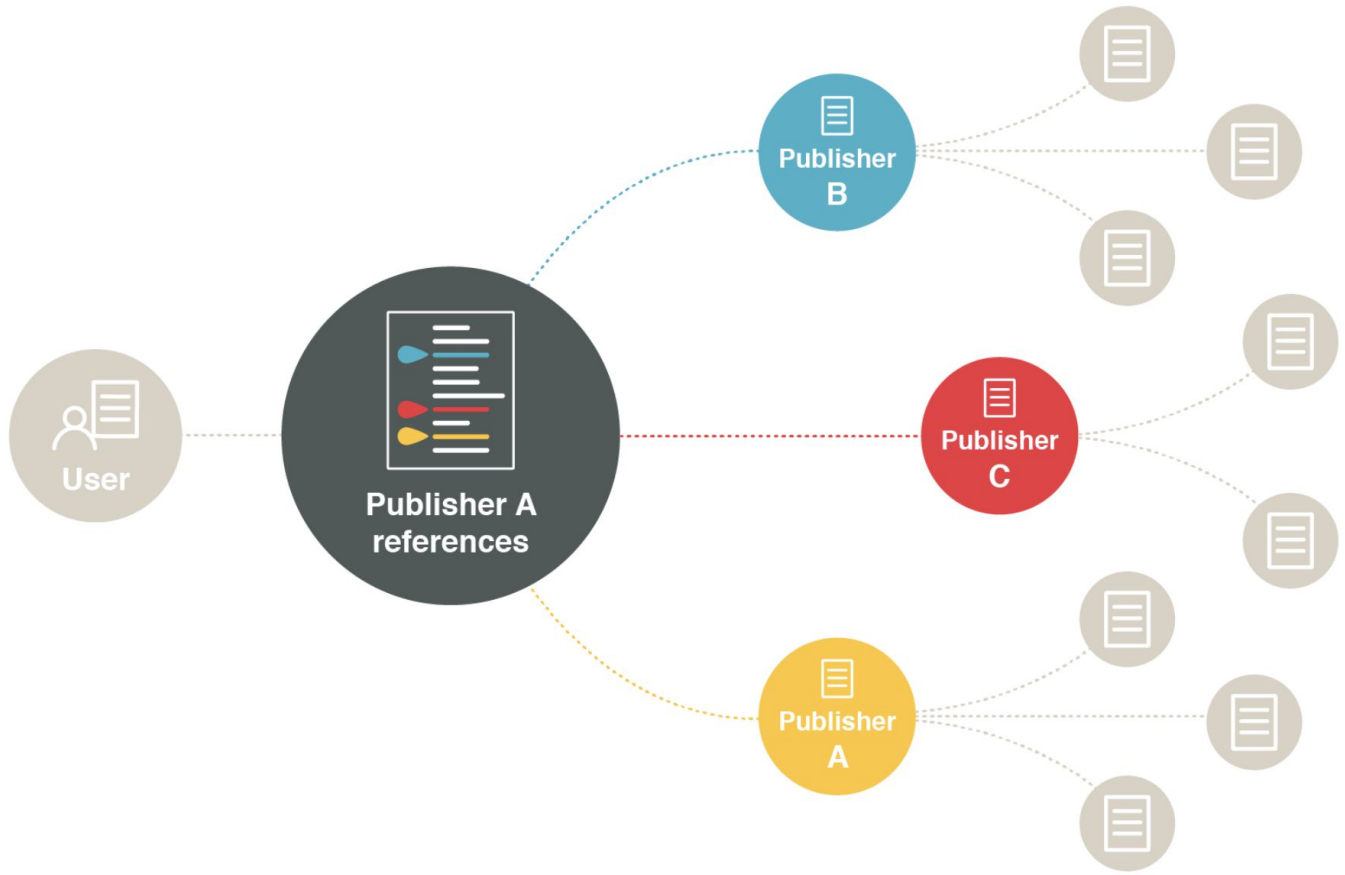


Crossref overview

- Over 12,000 member organizations
- 16 member board, cross section of international publishers
- Metadata store of over 113 million scholarly content items
- A DOI is just the start - we offer a wide array of services to ensure that scholarly research metadata is registered, linked, and distributed.
- We preserve the metadata we receive and make it available via our open APIs and search interfaces.

DOIs and metadata for research outputs







Gives us scope to support things like:

Helping researchers identify content they can text mine



Geoffrey Bilder – 2020 April 16
In [Metadata](#), [Community](#), [APIs](#)

TL;DR

Many organizations are doing what they can to aid in the response to the COVID-19 pandemic, for researchers to identify, locate, and access content for text mining. In order to do this, metadata that:

- Point to the full text of the content.
- Indicate that the content is available under an open access license or that it is being made

<https://www.crossref.org/blog/helping-researchers-identify-content-they-can-text-mine/>

I40C

[About](#) [Goals](#) [Publishers](#) [Stakeholders](#) [Founders](#) [FAQ](#) [News](#) [Press](#)

About

Citations are the links that knit together our scientific and cultural knowledge. They are primary data that provide both provenance and an explanation for how we know facts. They allow us to attribute and credit scientific contributions, and they enable the evaluation of research and its impacts. In sum, citations are the most important vehicle for the discovery, dissemination, and evaluation of all scholarly knowledge.

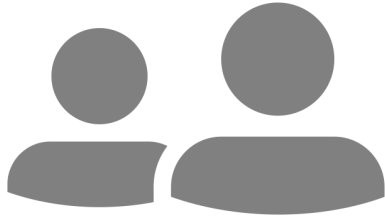
As the number of scholarly publications is estimated to **double** every nine years, citations – and the computational systems that track them – enable researchers and the public to keep abreast of significant developments in any given field. For this to be possible, it is essential to have unrestricted access to bibliographic and citation data in machine-readable form.

The present scholarly communication system inadequately exposes the knowledge networks that already exist within our literature. **Citation data are not usually freely available to access, they are often subject to inconsistent, hard-to-parse licenses, and they are usually not machine-readable.**

<https://i4oc.org/>

+ other
individual
identifiers





Created by Ayub Irawan
from Noun Project

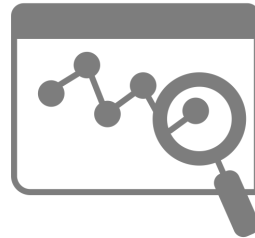


Created by Vincencio
from Noun Project

Created by Creative Stall
from Noun Project



Created by shashank singh
from Noun Project



Created by Millenials
from Noun Project



Created by shashank singh
from Noun Project

ALL
THE THINGS





https://commons.wikimedia.org/wiki/File:Ralls_Texas_Grain_Silos_2010.jpg
Leaflet / CC BY-SA (<https://creativecommons.org/licenses/by-sa/3.0>)

+ connections =





ORCID auto-update

🕒 3 minute read.

Crossref to Auto-Update ORCID Records



Ginny Hendricks – 2015 September 24
in [Member Briefing](#), [Metadata](#), [News Release](#), [News Release](#), [ORCID](#)

In the next few weeks, authors with an ORCID iD will be able to have Crossref automatically push information to their ORCID record. It's something that [ORCID users have been asking for](#) and we're pleased to see this integration. 230 publishers already include ORCID iDs in their metadata deposits with us, and currently 230 publishers already include ORCID iDs.

What this means for researchers



More visibility for your work! Crossref represents over 5000 scholarly publishers and

their ORCID iD and include it in the publication information they send us. Also it will mean less manual searching and adding, you've

ORCID Totals					
Statistic	October 2018	2018	April 2020	2020	% change 2018-2020
Total orcid/article pairs	3,059,399		6,627,744		117%
Total unique orcids	1,281,355		2,336,447		82%
Total works pushed	1,343,663		3,407,291		154%
Authors granting	502,822	39%	1,053,870	45%	110%
Authors denying	408,273	32%	563,967	24%	38%
Authors not responding	381,084	29%	732,011	31%	92%



We can see things like this:

Preprints are go at Crossref!



Rachael Lammey - 2016 November 02
In [Persistence](#), [Preprints](#)

We're excited to say that we've finished the work on our infrastructure to allow members to register preprints. We're doing this? Jennifer Lin [explains the rationale in detail](#) in an earlier post, but in short we want to help

- links to these publications persist over time
- they are connected to the full history of the shared research results
- the citation record is clear and up-to-date

Doing so will help fully integrate preprint publications into the formal scholarly record.

```
← → ↻ 🔒 api.crossref.org/works?filter=relation.type:has-preprint&facet=publisher-name:*&rows=0
Apps 26 CrossRef - Calend... SugarCRM - Com... CrossRef Administr... crossref.org The Secret Lif

{
  "status": "ok",
  "message-type": "work-list",
  "message-version": "1.0.0",
  "message": {
    "facets": {
      "publisher-name": {
        "value-count": 543,
        "values": {
          "Copernicus GmbH": 29835,
          "Springer Science and Business Media LLC": 5734,
          "MDPI AG": 5522,
          "Elsevier BV": 4168,
          "JMIR Publications Inc.": 3704,
          "Public Library of Science (PLOS)": 3459,
          "Oxford University Press (OUP)": 2541,
          "Wiley": 2280,
          "PeerJ": 1390,
          "eLife Sciences Publications, Ltd": 1353,
          "Frontiers Media SA": 1005,
          "American Society for Microbiology": 794,
          "Genetics Society of America": 697,
          "Informa UK Limited": 538,

```

https://api.crossref.org/works?filter=relation.type:has-preprint&facet=publisher-name:*&rows=0



And this:



UK RESEARCH AND INNOVATION

- Medical Research Council +
- Engineering and Physical Sciences Research Council +
- Economic and Social Research Council
- Arts and Humanities Research Council
- Biotechnology and Biological Sciences Research Council
- Research England
- Innovate UK
- Natural Environment Research Council +
- Science and Technology Facilities Council +

TYPE

- Journal Article (102,019)
- Conference Paper (990)
- Dataset (581)
- Posted Content (455)
- Chapter (113)
- Book (7)
- Report (1)

YEAR

- 2019 (21,856)

Metadata Search  Sign in

SORT BY: **RELEVANCE** PUBLICATION YEAR [DOWNLOAD AS CSV](#) PAGE 1 OF 104,166 RESULTS

Response of Jupiter's auroras to conditions in the interplanetary medium as measured by the Hubble Space Telescope and Juno

Journal Article published 16 Aug 2017 in Geophysical Research Letters volume 44 issue 15 on pages 7643 to 7652

Research funded by Science and Technology Facilities Council (ST/M005534/1ST/1004084/1HST-GO-14105.002-AST/K001000/1)

Authors: J. D. Nichols, S. V. Badman, F. Bagenal, S. J. Bolton, B. Bonfond, E. J. Bunce, J. T. Clarke, J. E. P. Connerney, S. W. H. Cowley, R. W. Ebert, M. Fujimoto, J.-C. Gérard, G. R. Gladstone, D. Grodent, T. Kimura, W. S. Kurth, B. H. M. ...

<https://doi.org/10.1002/2017gl073029> Actions

Spaceborne GNSS-Reflectometry on TechDemoSat-1: Early Mission Operations and Exploitation

Journal Article published Oct 2016 in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing volume 9 issue 10 on pages 4525 to 4539

Research funded by European Space Agency (4000109726) | ESA | NERC | Previous support for instrument development | Centre for Earth Observation and Instrumentation | InnovateUK | SEEDA | UKSA | Surrey Satellite Technology Ltd

Authors: Martin Unwin, Philip Jales, Jason Tye, Christine Gommenginger, Giuseppe Foti, Josep Rosello

<https://doi.org/10.1109/jstars.2016.2603846> Actions

Metal enhanced fluorescence biosensing: from ultra-violet towards second near-infrared window

Journal Article published 2018 in Nanoscale volume 10 issue 45 on pages 20914 to 20929

Research funded by Engineering and Physical Sciences Research Council (EP/R00661X/1)

Authors: Sarah Madeline Fothergill, Caoimhe Joyce, Fang Xie

<https://doi.org/10.1039/c8nr06156d> Actions



And this:

DataCite Blog

Datacite Citation Display: Unlocking Data Citations

January 7, 2020 by Kristian Garza · 7 min read

<https://doi.org/10.5438/1843-k679>

Are published datasets actually being cited? It's a question that must cross the mind of any avid user of this blog at least once. We have talked about data citation plenty on this blog. Often, though, we've talked about the links between academic resources and datasets — the kind of links collected by our [Event](#) tool. Many of these links between articles and datasets can be regarded as data citations, but not all.

Until today, it wasn't straightforward to visualize these links without having some programming background. Today, that changes. Today we introduce the Citations Display in DataCite Search. Here you will be able to access all these links and distinguish between data citations, references, and other relations that exist between academic resources with DOIs. In this blog post, we try to answer the questions you may have about DataCite & data citation.

Where does DataCite get information about the links between articles and datasets?4

There are two ways in which we get this information. DataCite members provide information about the links between their datasets and other academic resources. They do this by including related identifiers in their dataset DOI metadata they deposit with us. You can see the workflow in [Figure 1](#). Through our collaboration with Crossref, we also have information about dataset DOIs included in published articles. Like Crossref members, Crossref members include this information in the metadata they deposit with Crossref.

<https://blog.datacite.org/data-citation-display/>

7 citations No usage information was reported.

<https://doi.org/10.5438/0014> Cite Add to ORCID record

7 Citations

7 citations reported since publication in 2017.




7 Citations 2 References 2 Relations

- According to **Imperial College London** via **Datacite** the following collection **references** the item on this page: Rzepa, H. (2018). A FAIR to remember - The benefits of findable, accessible, interoperable and re-useable data. Imperial College London. <https://doi.org/10.14469/HPC/3657>
- According to **Imperial College London** via **Datacite** the following collection **references** the item on this page: Rzepa, H. (2018). Metadata Recommendations for DataCite Registration. Imperial College London. <https://doi.org/10.14469/HPC/4739>
- According to **GESIS Leibniz Institute For The Social Sciences** via **Datacite** the following publication **references** the




Gives scope to do things like this

 **Blog**

Home > Blog > **Free Public Data File of 112+ Million Crossref Records**

🕒 2 minute read.

Free Public Data File of 112+ Million Crossref Records

 [Jennifer Kemp](#) – 2020 April 09
In [Metadata](#), [Community](#), [APIs](#)

A lot of people have been using our public, open APIs to collect data that might be related to COVID-19. This is great and we encourage it. We also want to make it easier. To that end we have made a free data file of the public elements from Crossref's 112.5 million metadata records.

The file (65GB, in JSON format) is available via Academic Torrents here: <https://doi.org/10.13003/83B2GP>

It is important to note that [Crossref metadata](#) is always openly available. The difference here is that we've done the time-saving work of putting all of the records registered through March 2020 into one file for download.

<https://www.crossref.org/blog/free-public-data-file-of-112-million-crossref-records/>



But there's still work to do

Missing affiliation data #4 N

Closed jure opened this issue on 29 Sep 2014 · 11 comments

jure commented on 29 Sep 2014

Not sure if this is an actual issue, i.e., is this data available somewhere and is not exposed the API, or does CrossRef not collect it at all?

E.g. affiliations on a PLOS paper (http://api.plos.org/search?wt=json&start=0&rows=10&fl=id,title_display,subject,counter_total_month,affiliate_author_display,abstract,journal_article_type,publication_date&q=article_type_facet:%22issue%20image%22%20AND%20doc_type:full&sort=counter_total_month%20desc&api_key=7Jne3TIPu6dQFCK&=title:%22hello%22):

```

"affiliate": [
  "School of Psychology, College of Science and Engineering, University of Glasgow, Glasgow, United Kingdom",
  "Department of Psychology, Princeton University, Princeton, New Jersey, United States of America",
  "Voice Neurocognition Laboratory, Institute of Neuroscience and Psychology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, Scotland, United Kingdom",
  "Département de Psychologie, Université de Montréal, Montréal, Québec, Canada",
  "Institut des Neurosciences de La Timone, Université Aix-Marseille, Marseille, France",
  "Northwestern University, United States of America"
]

```

And CrossRef doesn't have it: <http://api.crossref.org/works/10.1371/journal.pone.0090779>

ckoscher commented on 29 Sep 2014

CrossRef's deposit schema allows for the inclusion of author affiliation. However, PLOS is not sending that data into us.

Chuck

Assignees
No one assigned

Labels
enhancement
reindex

Projects
None yet

Milestone
No milestone

Linked pull requests
Successfully merging a pull request may close this issue.
None yet

5 participants



<https://ror.org/00hswnk62>

Queen's University Belfast
QUB, PRIFYSGOL Y FRENHINES, BELFFAST

WEBSITE
<http://www.qub.ac.uk/>

OTHER IDENTIFIERS
 GRID [grid.4777.3](https://orcid.org/0000-0001-9100-0000)
 ISNI [0000000403747521](https://orcid.org/0000-0000-4037-47521)
 Crossref Funder ID [501100000873](https://orcid.org/501100000873)
 Wikidata [Q877925](https://orcid.org/Q877925)

UNITED KINGDOM EDUCATION

<https://ror.org>



Grant IDs

“While most funders state that the majority of their reporting requests are fulfilled, much of the information reported is provided late or of low quality and requires time-consuming cleanup.”



Created by Neil Schmitzer
from Neat Project

The screenshot shows the Europe PMC website interface. At the top, there is a navigation bar with 'Europe PMC' logo and links for 'About', 'Tools', 'Developers', 'Help', and 'Europe PMC plus'. Below this is a search bar with the text 'Search worldwide, life-sciences literature' and a search button. A navigation menu below the search bar includes 'Tools overview', 'ORCID article claiming', 'Journal list', 'Grant finder' (highlighted), 'External links service', 'RSS feeds', 'Annotations', and 'Annotations submission service'. The main content area displays the following information:

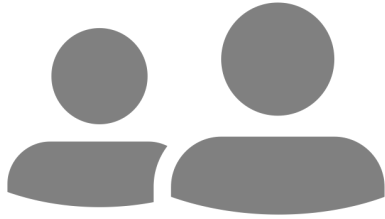
- Initiative to Develop African Research Leaders (IDEAL)**
- Dr SM Kinyanjui | Kenya Medical Research Institute (Kemri)
- Abstract**
- Research is key in tackling the health challenges that Africa faces. In KWTRP we have been committed to building sustainable capacity alongside an active and diverse research programme covering social science, health services research, epidemiology, laboratory science including molecular biology and bioinformatics. Our strategy has been successful in delivering high quality PhD training, leveraging individual funding and programme funding in order to place students in productive groups and provide high quality supervision and mentorship. Here we plan to consolidate and build on these outputs to address long-term sustainability. We will emphasise the full career path needed to generate research leaders. KWTRP aims to address capacity building for research through an initiative that employs a progressive and long term outlook in the development of local research leadership. The overall aim of the "Initiative to Develop African Research Leaders" (IDEAL) is to build a critical mass of African researchers who are technically proficient as scientists and well-equipped to independently lead science at international level, able to engage with funders, policy makers and governments, and to act as supervisors and mentors for the next generation of researchers.
- Funded by** Wellcome Trust
- £ 8,037,752**
- Duration** 01 Jan 2016 - 31 Dec 2020
- Internal grant ID** 107769
- Grant DOI** 10.35802/107769 (highlighted with a red box)
- Funding stream** Population and Public Health
- Grant type** DELTAS Initiative



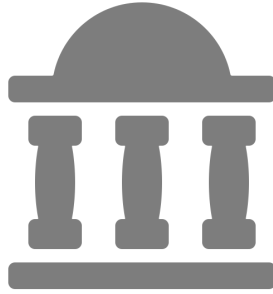
HT Elizabeth Gadd

The next platform

1. **Preprints** are available for immediate review and consumption by all (we know that in many disciplines [peer review doesn't materially change the output](#) so this is a must – and it shaves years off publication times).
2. Post-publication **peer review reports** give reviewers credit.
3. Approved outputs can be **indexed in bibliographic databases** and made just as discoverable as journal-based outputs.
4. All outputs can be made available under a **CC-BY licence** and in accordance with all the necessary technical requirements for truly findable, accessible, interoperable and reproducible research.
5. **Links to datasets and other open research outputs can be added.**
6. UKRI gets to trace, cradle-to-grave, the **impacts of their funded research** as it's all published in one place.
7. Publishing on the platform doesn't depend on the wealth of the organisation and the size of their Gold OA Fund.
8. It reduces researcher anxiety about 'getting stuff published'. If it's funded by UKRI it is welcome on the platform.
9. Quality is assessed through peer review reports, and impact through **subsequent usage (citations if that floats your boat)**, and not by journal brand.



Created by Ayub Irawan
from Noun Project

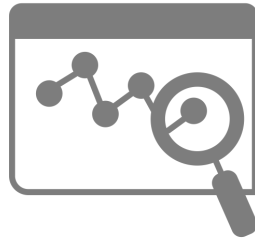


Created by Vincencio
from Noun Project

Created by Creative Stall
from Noun Project



Created by shashank singh
from Noun Project



Created by Millenials
from Noun Project



Created by shashank singh
from Noun Project



Thank you!
Questions?

rlammey@crossref.org

<https://crossref.org>