



Crossref

Using Crossref metadata to track Open Access (and more)



Rachael Lammey
@CrossrefOrg

OpenAIRE Workshop, Oslo, Feb 2017



We are Crossref, a not-for-profit membership organization for scholarly publishing, working to make content easy to find, cite, link and assess. We do it in five ways: rallying the community; tagging metadata; running a shared infrastructure; playing with new technology; and making tools and services to improve research communications.

It's as simple—and as complicated—as that.



Crossref Overview

- 5500+ publisher members, 1 member 1 vote
- Metadata store of over 85 million scholarly content items (including preprints)
- Persistent citation linking
- Funder identifiers
- Report and display corrections & retractions
- Check manuscripts for similarities
- Data about content activity e.g. social, dataset links
- Open REST API & Search

infrastructure |'ɪnfəstrʌktʃə|

noun

the basic physical and organizational structures and facilities (e.g. buildings, roads, power supplies) needed for the operation of a society or enterprise. The social and economic infrastructure of a country.

<metadata>

Why do publishers assign DOIs?

Show people where your content is located, and update that if/when the content moves

- Helps indexing
- Linking from other sources is persistent
- Full-text links can be used for text-mining and indexing



More metadata means that people have more ways to find publications

What are we trying to do?

- Institutions want to be able to track publications from their researchers and how much content is being published under OA licenses
- Funders need a way to track outputs from thousands of publishers
- Publishers will need a standard and efficient way to demonstrate conformance mandates
- All the stakeholders in the process (funders, publishers, institutions and researchers) will span disciplines, institutions, geographies and jurisdictions

How can we do it?

- Crossref has extended its metadata schemas and Application Programming Interfaces (APIs) to enable institutions, funders and publishers to use Crossref metadata to:
 - track research that is subject to these mandates
 - ensure that said research is being disseminated according to the requirements of the mandates

Metadata can be used to identify

- Publications relating to research supported by particular funders.
- Publications from particular researchers identified by their ORCID ID
- Bibliographic metadata
- License information (including embargoes)
- The location of the full text for reading and text & data mining application
- Long-term preservation arrangements
- Clinical trial information
- Updates (errata, corrigenda, retractions)

Funding and license information

- Standardisation of funder names using the Open Funder Registry
- 14,200 international funders
- Publisher deposit of funding metadata which they collect from authors upon submission (or add it afterwards)
- Facilitates large-scale analysis and reporting to funders

Made available in bulk or via an API

```
← → ↻ ⓘ api.crossref.org/works?filter=has-funder:true&rows=0

{
  "status": "ok",
  "message-type": "work",
  "message-version": "1",
  "message": {
    "facets": {},
    "total-results": 1,
    "items": [],
    "items-per-page": 10,
    "query": {
      "start-index": 0,
      "search-terms": ""
    }
  }
}

{
  "date-time": "2015-11-13T10:03:35Z",
  "timestamp": 1447409015813,
  "reference-count": 0,
  "publisher": "Springer Science + Business Media",
  "funder": [
    {
      "award": [ ],
      "name": "Arthritis Research UK",
      "DOI": "10.13039/501100000341"
    }
  ],
  "content-domain": {
    "domain": [ ],
    "crossref-id": "10.1007/978-1-4939-9736-7_1"
  }
}
```

Search 14,200 funders connected to
1,252,198 published works with funding
data

 Search funders...

<http://search.crossref.org/funding>

RESEARCH COUNCILS UK

Arts and Humanities Research Council

Biotechnology and Biological Sciences
Research Council

Natural Environment Research Council +

Science and Technology Facilities
Council

Medical Research Council

Engineering and Physical Sciences
Research Council

Economic and Social Research Council

TYPE

- ☐ Journal Article (5,234)
- ☐ Dataset (449)
- ☐ Chapter (12)
- ☐ Conference Paper (2)
- ☐ Book (1)

YEAR

SORT BY: RELEVANCE PUBLICATION YEAR

DOWNLOAD AS CSV

PAGE 1 OF 5,698 RESULTS

Beyond the Medial Regions of Prefrontal Cortex in the Regulation of Fear and Anxiety

Journal Article published 22 Feb 2016 in **Frontiers in Systems Neuroscience** volume 10

Research funded by Medical Research Council (G0901884)

Authors: Yoshiro Shiba, Andrea M. Santangelo, Angela C. Roberts

<https://doi.org/10.3389/fnsys.2016.00012> Actions

EMOTICOM: A Neuropsychological Test Battery to Evaluate Emotion, Motivation, Impulsivity, and Social Cognition

Journal Article published 24 Feb 2016 in **Frontiers in Behavioral Neuroscience** volume 10

Research funded by Medical Research Council (MR/J011894/1)

Authors: Amy R. Bland, Jonathan P. Roiser, Mitul A. Mehta, Thea Schei, Heather Boland, Daniel K. Campbell-Meiklejohn, Richard A. Emsley, Marcus R. Munafo, Ian S. Penton-Voak, Ana Seara-Cardoso, Essi Viding, Valerie Voon, Barbara J. Sahakian ...

<https://doi.org/10.3389/fnbeh.2016.00025> Actions

Sepsis Patients with First and Second-Hit Infections Show Different Outcomes Depending on the Causative Organism

Journal Article published 26 Feb 2016 in **Frontiers in Microbiology** volume 7

License information

- Information available as an overview or at the work level

A screenshot of a web browser window. The address bar shows the URL: `api.crossref.org/works?rows=0&filter=license.url:http://creativecommons.org/licenses/by/3.0/`. The browser interface includes back, forward, and refresh buttons.

```
{
  status: "ok",
  message-type: "work-list",
  message-version: "1.0.0",
  - message: {
    - query: {
      search-terms: null,
      start-index: 0
    },
    items-per-page: 0,
    items: [ ],
    total-results: 216394,
    facets: { }
  }
}
```

At the work level

```

<!--
▼<publisher_item>
  <identifier id_type="pii">6742786</identifier>
  <identifier id_type="other">6742786</identifier>
</publisher_item>
▼<program name="AccessIndicators">
  <license_ref>http://creativecommons.org/licenses/by/4.0/</license_ref>
</program>
▼<doi_data>
  <doi>10.1155/2016/6742786</doi>
  <timestamp>20161031090007</timestamp>
  <resource>https://www.hindawi.com/journals/ae/2016/6742786/</resource>
▼<collection property="text-mining" setbyID="hindawi">
  ▼<item>
    ▼<resource mime_type="application/pdf">
      http://downloads.hindawi.com/journals/ae/2016/6742786.pdf
    </resource>
  </item>
</collection>
-->

```

License information

- Can convey embargo & archiving information:

```
</fr:assertion>
</fr:program>
<ai:program name="AccessIndicators">
  <ai:license_ref applies_to="vor" start_date="2014-05-02">
    http://publisher.com/licenses/license_v1.html
  </ai:license_ref>
  <ai:license_ref applies_to="am">
    http://publisher.com/licenses/open_license_v1.html
  </ai:license_ref>
  <ai:license_ref applies_to="tdm" start_date="2014-12-02">
    http://publisher.com/licenses/tdm_license_v1.html
  </ai:license_ref>
</ai:program>
<archive_locations>
  <archive name="CLOCKSS"/>
</archive_locations>
```

ORCID iDs

- Nearly 0.6 million ORCID iDs deposited with Crossref from nearly 850 publishers
- Enables auto-update workflow
- Enables accurate mapping between authors and publications

```
    },  
    source: "CrossRef",  
-   title: [  
        "الإعجاز في قوله تعالى والمرسلات عرفا"  
    ],  
    prefix: "http://id.crossref.org/prefix/10.19138",  
-   author: [  
        - {  
            affiliation: [ ],  
            family: "معوض",  
            given: "رامي",  
            ORCID: "http://orcid.org/0000-0003-4211-3903"  
        }  
    ]  
}
```

Measuring impact

- Citations

- Publishers can make these openly available via Crossref

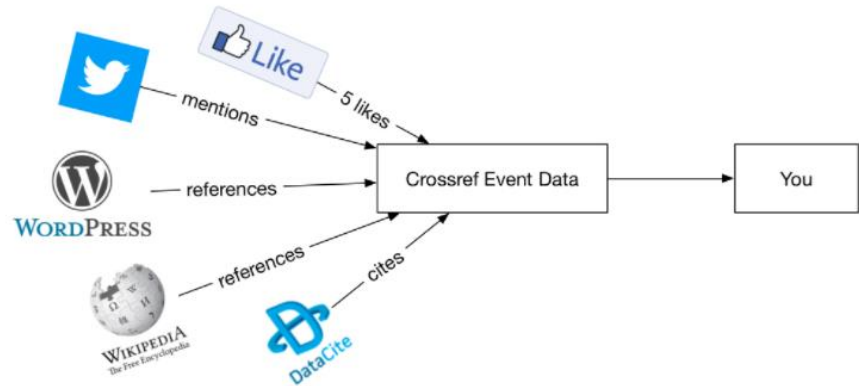
- Crossref Event Data

- Our data is open, auditable, transferrable, and portable
 - Regular updates from data contributors and reliable, secure storage
 - Raw data, free of charge with the addition of an SLA provided at favorable rates
 - A community-led initiative and data from a neutral source
 - <https://www.crossref.org/services/event-data/>

Crossref Event Data

Every 'thing that happens' is recorded as an individual Event. We gather Events from a wide range of sources, but examples include:

- an article was linked from DataCite dataset via its Crossref DOI
- an article was referenced in Wikipedia using its Crossref DOI
- an article was mentioned on Twitter using its Article Landing Page URL
- an article has been liked on Facebook 55 times as of June 15th



Also:

- Free to read tag
- Work on organisational identifiers (the OI project):
<http://blog.crossref.org/2016/10/the-oi-project.html>
- Early content registration slated for
Q3:<http://blog.crossref.org/2016/03/community-responses-to-our-proposal-for-early-content-registration.html>

The Crossref extended APIs, of course, will only work if publishers supply the appropriate metadata and follow the specified metadata deposit workflows.

Thank you

@CrossrefOrg
rlammey@crossref.org
<https://crossref.org>

