

Making the most of Crossref metadata



Deutsches Zentrum für Hochschul- und
Wissenschaftsforschung GmbH (DZHW)

Jennifer Kemp, Head of Partnerships, Crossref



Our mission

Crossref makes research outputs easy to find, cite, link, assess, and reuse.

We're a not-for-profit membership organization that exists to make scholarly communications better.

Why the metadata is so important

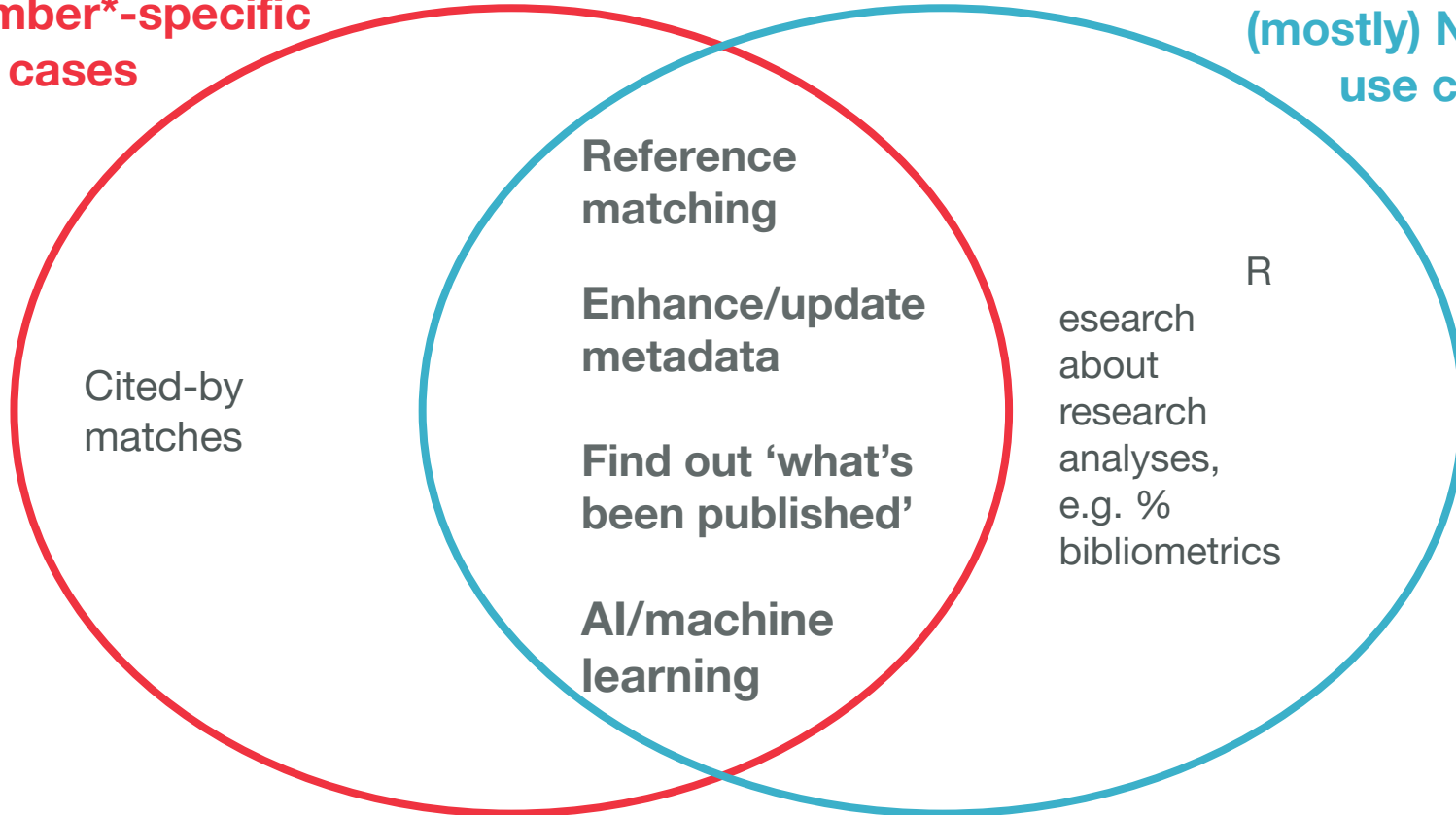
600+ million
queries per month



High-Level Use Cases

**Member*-specific
use cases**

**(mostly) Non-member
use cases**



**May be done by third parties, e.g. hosting platforms*

Strategic goals 2025



Bolster the team



Improve our metadata



Live up to POSI



Collaborate & partner

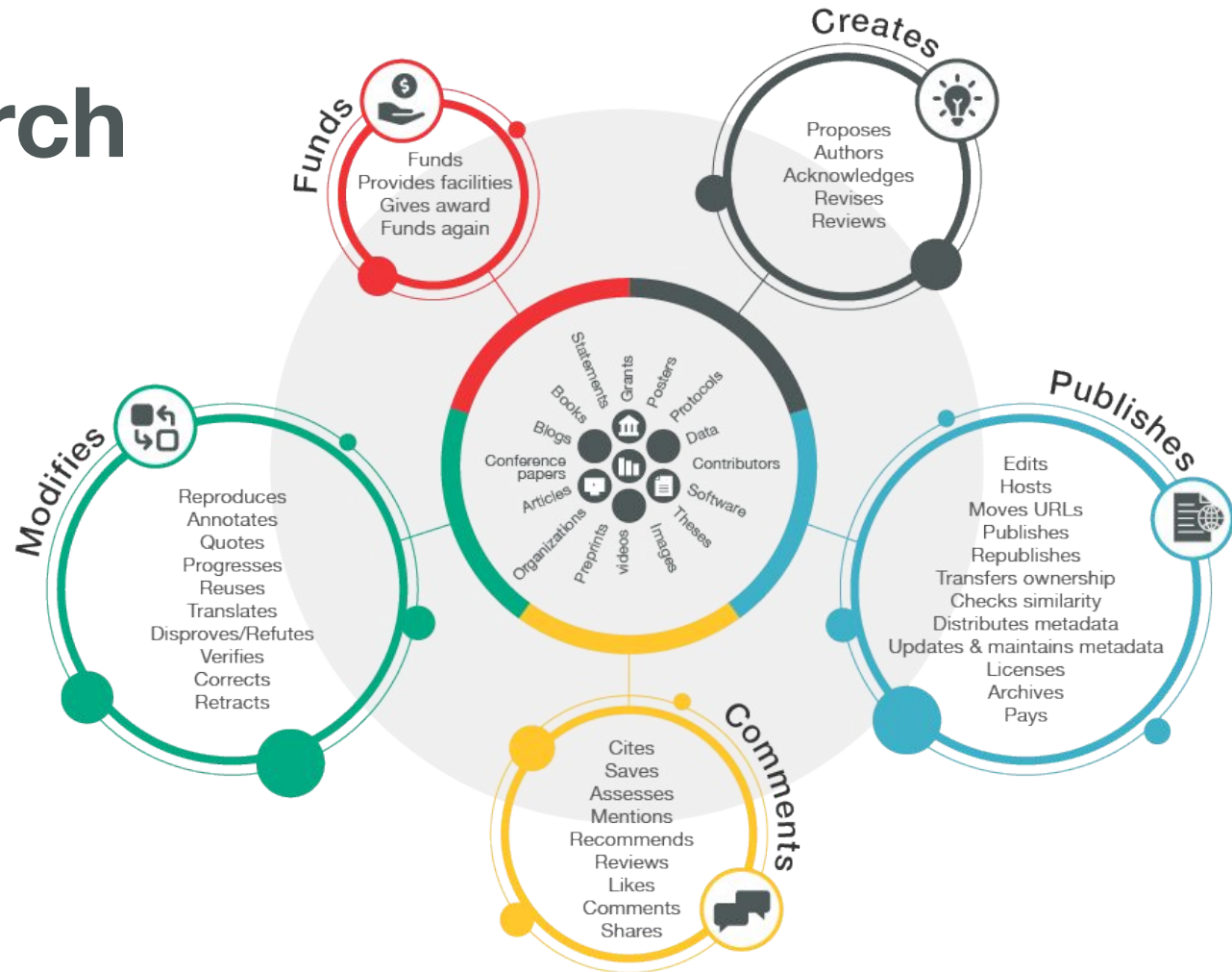


Engage communities



Simplify & enrich services

The research nexus



What's in the metadata (& what's not)

reference-count: 33
publisher: "Public Library of Science (PLOS)"
issue: "2"

▼ license:

▼ 0:

▼ start:

▼ date-parts:

▼ 0:

0: 2022

1: 2

2: 17

date-time: "2022-02-17T00:00:00Z"

timestamp: 1645056000000

content-version: "vor"

delay-in-days: 0

URL: "http://creativecommons.org/licenses/by/4.0/"

▼ funder:

▼ 0:

DOI: "10.13039/100009941"

name: "Ministry of Defence"

doi-asserted-by: "publisher"

▼ content-domain:

▼ domain:

0: "www.plosone.org"

crossmark-restriction: false

▼ short-container-title:

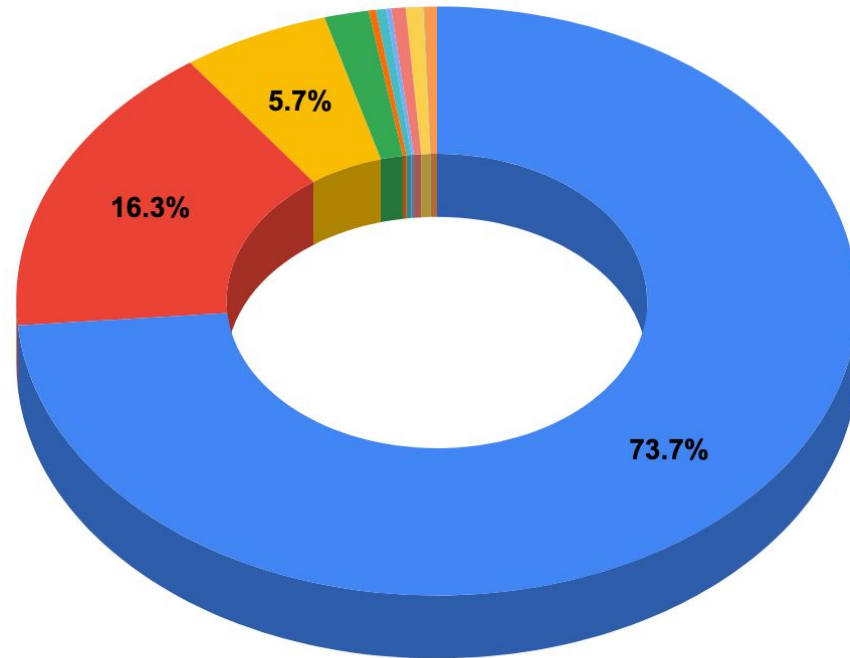
0: "PLoS ONE"

▼ abstract:

"<jats:sec id=\"sec001\">\n<jats:title>Background</jats:title>\n<jats:p>Medical personnel may find it challenging to distinguish severe Exertional Heat Illness (EHI) with attendant risks of organ-injury and longer-term sequelae, from lesser forms of incapacity associated with strenuous physical exertion. Early evidence for injury at point-of-incapacity could aid the development and application of targeted interventions to improve outcomes. We aimed to investigate whether biomarker surrogates for end-organ damage sampled at point-of-care (POC) could discriminate EHI versus successful marathon performance.</jats:p>\n</jats:sec>\n<jats:sec id=\"sec002

Records by type

- Journals
- Books
- Conferences
- Datasets
- Standards
- Dissertations
- Peer review reports
- Reports
- Posted content/preprints
- Grants
- Other



← Example grant record

Summary indicator of all records



Crossref Stats Page: Updated on Mar 15, 2022

MetaData Deposit

Total Records in Database	133,634,609
Total Prefixes	23,014
Number of Journals	101,836
Number of Books	1,700,274
Number of Conference Proceedings	87,764
Number of Journal DOIs	94,378,851
Number of Book DOIs	21,374,459
Number of Conference DOIs	7,296,735
Number of Components	5,512,894
Number of Standards Titles	436,095
Number of Standards DOIs	346,663
Number of Databases	58,769
Number of Database DOIs	2,272,905
Number of Articles with References Deposited	55,335,629
Number of Cited-By Links	1,301,969,744
Number of DOIs with at least one cited-by link (UNDER CONSTRUCTION)	0
Number of Journal Title Level DOIs	47,345
Number of Book Title Level DOIs	1,722,742
Number of ConfProc Title Level DOIs	55,539
Number of Dissertation Title Level DOIs	505,179
Number of Report Title Level DOIs	706,879
Number of Standards Title Level DOIs	346,098
Number of Database Title Level DOIs	42,929
Number of Unique DOIs with FundRef data	2,947,660

```
"publisher": "Wellcome",
"award": "107769",
"DOI": "10.35802/107769",
"type": "grant",
"created": {
"date-parts": [
[
2019,
9,
25
]
],
"date-time": "2019-09-25T07:17:20Z",
"timestamp": 1569395840000
},
"source": "Crossref",
"prefix": "10.35802",
"member": "13928",
"project": [
{
"project-title": [
{
"title": "Initiative to Develop African Research Leaders (IDeAL)"
}
],
"project-description": [
{
"description": "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.",
"language": "en"
}
],
}
```

Participation Reports

https://www.crossref.org/members/rep/

Participation Reports

How good is your metadata?

Find a member

Search by member

Learn more

Richer metadata makes content useful.
Make sure your work can be found.

Find a member

Learn more

Richer metadata makes content useful.
Make sure your work can be found.

Institute of Electrical and Electronics Engineers (IEEE)

582,413
Total registered content items

Content type: Journal articles

Preprints 7,224 Journal articles 195,569 Conference papers 379,620

Journal articles

Search by title

Current content

References



Open references



ORCID IDs



Funder Registry IDs



Funding award numbers



Crossmark enabled



Text mining URLs



License URLs



Similarity Check URLs



Abstracts



Screenshot

Conference Proceeding metadata

Home > Documentation > Content Registratio

Documentation

The importance of metadata

Setting up as a member

Content Registration

Descriptive metadata

Administrative metadata

Structural metadata

Introduction to content types

Content type markup guide

Required, recommended, and optional elements

Screenshot

Required

Series	titles, ISSN
Proceeding level	proceedings_title, publisher, publication_date (year)
Conference paper	contributors , titles, doi_data

Recommended

Series level	doi_data, contributors , series_number, ISBN
Conference level	volume, contributors , ISBN, event_metadata (conference_date, conference_location, conference_acronym, conference_theme, conference_sponsor, conference_number), proceedings_subject
Conference paper	publication_date, pages, citation_list, funding , license , and Crossmark metadata

Optional

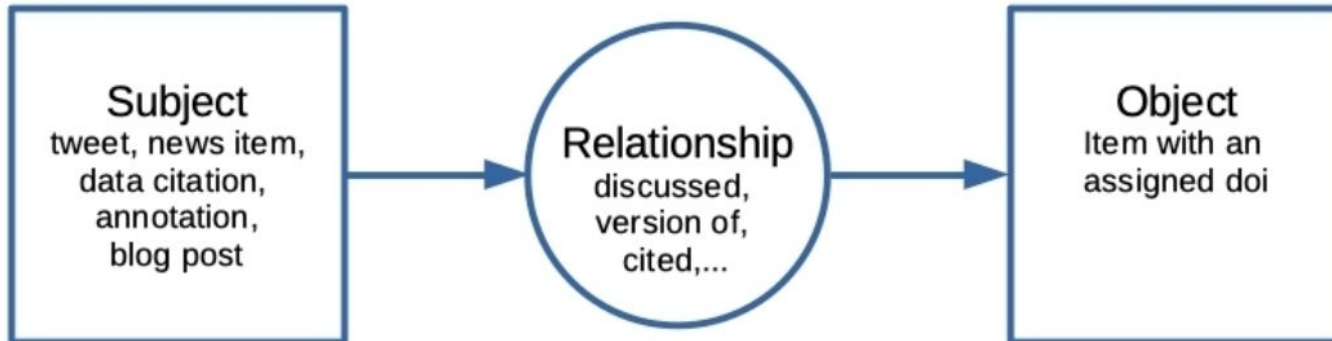
[publisher_item](#), [coden](#), [component_list](#)

Event Data & relationships



Home > Find a service > **Event Data**

When someone links their data online, or mentions research on a social media site, we capture that event and make it available for anyone to use in their own way. We provide the unprocessed data—you decide how to use it.



How to get the metadata



Search the metadata of journal articles, books, standards, datasets & more

Can Richer Metadata Rescue Research?

Journal Article published 31 May 2018 in *The Serials Librarian* volume 74 issue 1-4 on pages 207 to 211

Authors: Jennifer Kemp, Clare Dean, John Chodacki

<https://doi.org/10.1080/0361526x.2018.1428483> Actions

Metadata and discoverability: A use

Journal Article published 25 Oct 2018 in *Information Resources for Learning* volume 38 issue 1-2 on pages 81 to 84

Authors: Jennifer Kemp

Editors: Bonnie Lawlor

<https://doi.org/10.3233/isu-180004> Actions

- Cite
- Filter on this publication
- Metadata as JSON

Feature / option	Metadata Search	Simple Text Query	REST API	XML API	OAI-PMH	OpenURL	Metadata Plus (OAI-PMH + REST API)
Interface for people or machines?	People	People	People (low volume and occasional use) and machines	Machines	Machines	Machines	Machines
Output format	Text, JSON	Text	JSON	XML	XML	XML	JSON, XML
Suitable for citation matching?	Yes (low volume)	Yes	Yes	Yes	No	No	Yes
Supports volume downloads?	No	No	Yes	No	Yes	No	Yes
Suitable for usage type	Frequent and occasional	Frequent and occasional	Frequent and occasional	Frequent	Frequent	Frequent	Frequent and occasional
Free or cost?	Free	Free	Free and cost options	Free and cost options	Cost for full service, more options available	Free	Cost
Includes all available metadata?	In JSON only	DOIs only	Yes	Yes	Yes	Bibliographic only	Yes

REST API - lots of query options, e.g

This endpoint supports the following facets:

- `affiliation` - author affiliation
- `archive` - archive location
- `assertion` - custom Crossmark assertion name
- `assertion-group` - custom Crossmark assertion group name
- `category-name` - category name of work
- `container-title` - [max value 100], work container title, such as journal title, or book title

Sample

Being able to select random results is useful for both testing and sampling. You can use the `sample` parameter to retrieve random results. So, for example, the following selects 10 random works:

```
/works?sample=10
```

Note that when you use the `sample` parameter, the `rows` and `offset` parameters are ignored.

Parameter combinations

Any combination of `query`, `query.*`, `filter`, `facet`, `select` and `sort` can be used with offsets. Sampling cannot be combined with offsets.

Any combination of `query`, `query.*`, `filter`, `facet`, `select` and `sort` may also be used with deep paging cursors. `rows` may also be specified.

`offset` and `sample` cannot be used in combination with cursors.

 2 minute read.

New public data file: 120+ million metadata records



[Jennifer Kemp](#) – 2021 January 19

In [Metadata](#), [Community](#), [APIs](#)

2020 wasn't all bad. In April of last year, we released our [first public data file](#). Though Crossref metadata is always openly available—and our board recently cemented this by voting to adopt the [Principles of Open Scholarly Infrastructure \(POSI\)](#)—we've decided to release an updated file. This will provide a more efficient way to get such a large volume of records. The file (JSON records, 102.6GB) is [now available](#), with thanks once again to Academic Torrents.

Related options: ROR, DataCite, etc.

Find a ROR ID: ror.org/search



Search Registry...

 <https://ror.org/03yrm5c26>

California Digital Library

CDL

WEBSITE

<http://www.cdlib.org/>

OTHER IDENTIFIERS

GRID [grid.463323.3](#)

ISNI [0000000119575136](#)

Wikidata [Q5020447](#)

UNITED STATES

ARCHIVE

@ResearchOrgs



Thank-you

Questions?

jkemp@crossref.org

support@crossref.org

More information

<https://www.crossref.org/blog/>

<https://community.crossref.org/>

<https://www.crossref.org/documentation/>

<https://api.crossref.org/>



@CrossrefOrg