



Participation Reports help Crossref members drive research further

Anna Tolwinska

Crossref, Wakefield, MA, USA

Abstract

This article aims to explain the key metadata elements listed in Participation Reports, why it's important to check them regularly, and how Crossref members can improve their scores. Crossref members register a lot of metadata in Crossref. That metadata is machine-readable, standardized, and then shared across discovery services and author tools. This is important because richer metadata makes content more discoverable and useful to the scholarly community. It's not always easy to know what metadata Crossref members register in Crossref. This is why Crossref created an easy-to-use tool called Participation Reports to show editors, and researchers the key metadata elements Crossref members register to make their content more useful. The key metadata elements include references and whether they are set to open, ORCID iDs, funding information, Crossmark metadata, licenses, full-text URLs for text-mining, and Similarity Check indexing, as well as abstracts. ROR IDs (Research Organization Registry Identifiers), that identify institutions will be added in the future. This data was always available through the Crossref's REST API (Representational State Transfer Application Programming Interface) but is now visualized in Participation Reports. To improve scores, editors should encourage authors to submit ORCIDs in their manuscripts and publishers should register as much metadata as possible to help drive research further.

Keywords

Crossref; Metadata; Participation Reports; Research infrastructure; Scholarly communications

Introduction

Background: Metadata is heavily relied upon by both researchers and the scholarly community as it helps drive new discoveries. This became especially important during the coronavirus disease 2019 (COVID-19) pandemic. Sharing metadata openly helped researchers make important connections, build upon previous research, and perhaps played a part in the work to create the COVID-19 vaccine. As a scholarly infrastructure provider, Crossref played a role through collecting the metadata from its members (who include publishers and funders), storing it, and then distributing it in an open, standardized and machine-readable format to discovery services and other tools that researchers use worldwide. This is important because open scholarly in-

Received: July 1, 2021

Accepted: July 24, 2021

Correspondence to Anna Tolwinska
atolwinska@crossref.org

ORCID

Anna Tolwinska

<https://orcid.org/0000-0001-5088-8915>

Infrastructure has become critical to many in the research community, especially in light of some more commercially-run infrastructure being discontinued. In June of 2020 and again in January of 2021, Crossref released over 100 million metadata records as a large public data file to help spur the efforts of researchers. The combined power of all of our members' metadata enabled the community to use it in creative ways and build tools that help drive important discoveries.

Crossref collects a lot of metadata from its members but it is not always easy to see what metadata is in Crossref. Authors and editors want to see what metadata their production teams or vendors register with Crossref because richer metadata can help increase the discovery of their journals and publications. However, without querying Crossref's Representational State Transfer Application Programming Interface (REST API) it was really hard for them to get at this information in the past. So Crossref created an easy-to-use tool called Participation Reports, which helps publishers, editors, authors, and researchers see the most important or key metadata elements Crossref members are registering.

Objectives: This article aims to explain each key metadata element, why it's important to check the reports regularly, and how Crossref members can improve their scores. Specifically presented as follows: definition of participation reports, its importance in research, ten key elements of the reports, editor's role to improve the coverage in participation reports, and future plans.

What Are Participation Reports?

Participation Reports provide an easy way to see coverage of ten key metadata elements above and beyond the basic bibliographic metadata that all members are obliged to provide [1]. This includes metadata such as ORCID iDs for contributors, funding acknowledgements, reference lists, and abstracts—richer metadata that makes content more discoverable, and much more useful to the scholarly community as a whole, including among members themselves [2]. It's a visualization of the metadata that's already available via our public (REST API), except it's much easier to use. It is a place where anyone can see the metadata coverage of members, and members themselves can track their progress over time to see what's already registered and what's still missing. They are free and open to everyone and don't require a login. Participation Reports from the Korean Council of Science Editors was presented in Fig. 1.

Why Are Participation Reports Important?

Metadata helps move research forward and helps make new

discoveries happen. Crossref members that register a lot of metadata, especially richer metadata, help make content useful to researchers and the scholarly community. Members are not always aware however if they're registering the key metadata elements that help make those important connections and drive research further. Some rely on vendors or third parties to send their metadata deposits to Crossref and that makes it even harder to know exactly what metadata they are registering. Participation Reports show members and editors exactly what key metadata is being registered for their journals or publications, why it's important, what's still missing, and how to fill in the gaps.

What Metadata Does Crossref Collect?

Crossref collects a lot of metadata but not all of it is displayed in the Participation Reports. We have different types of metadata that our members register with us and that metadata serves many different purposes. We require basic bibliographic metadata to register a DOI, but it's the richer metadata that makes content go even further - for example, being able to find articles via an ORCID iD, who funded the research or the license it is published under all helps too.

Explaining Administrative, Descriptive, and Structural Metadata

Administrative metadata provides information about the origin and maintenance of a research object. This includes a link to accessing its full-text. Administrative metadata includes information needed to support the preservation of a research object, including archiving arrangements.

Descriptive (bibliographic) metadata consists of metadata used to describe and cite an item. Examples of bibliographic metadata include authors, titles, pages, dates. The bibliographic metadata registered with Crossref is used mainly in matching DOIs to citations and capturing citations in reference management tools.

The third type of metadata is structural metadata, which provides information about how research objects are organized, both within a research object (for example, a book composed of chapters, chapters composed of pages, and pages arranged in a particular order), and relationships between research objects (for example, a preprint, version of record, and a dataset).

What Metadata is Displayed in the Participation Reports?

All of the administrative, descriptive, and structural metadata

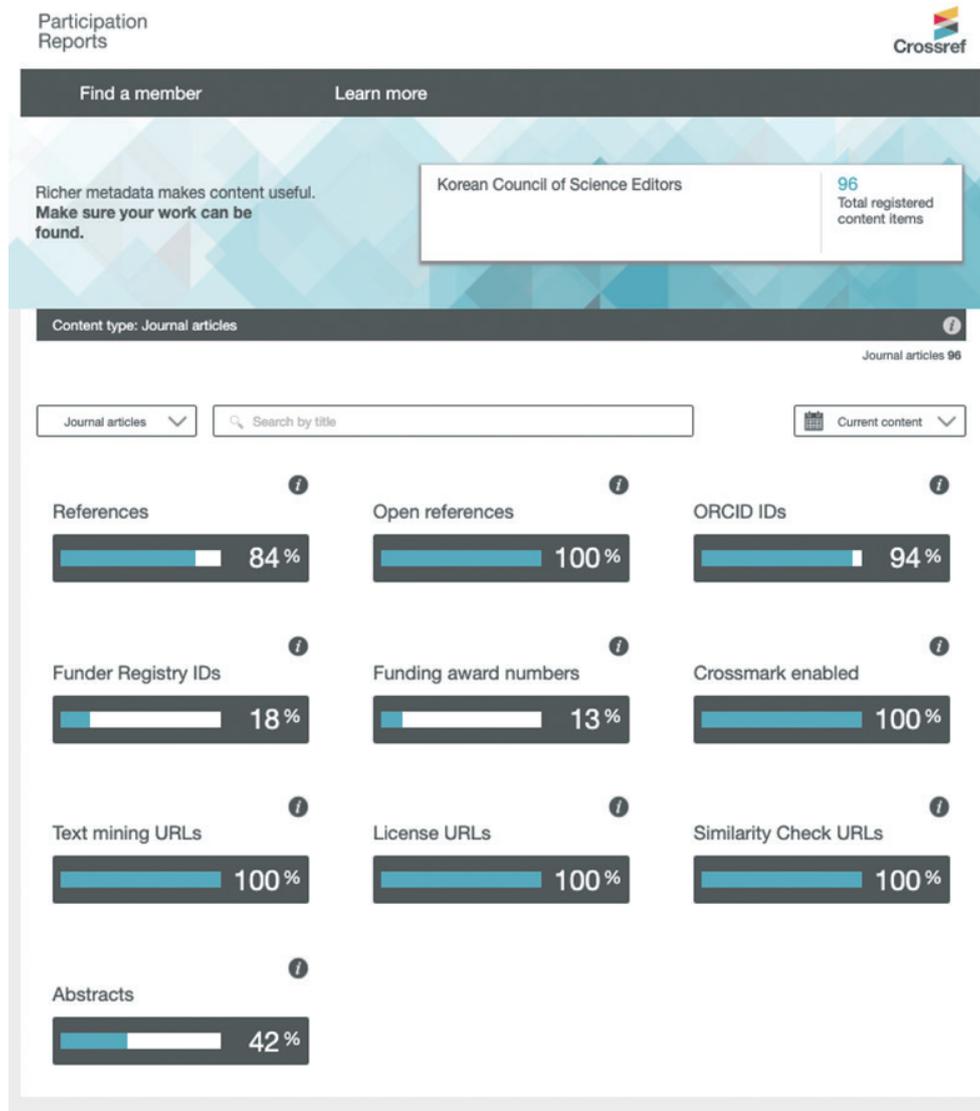


Fig. 1. Participation Reports example from the Korean Council of Science Editors.

is available via our REST API but in Participation Reports Crossref chose to display 10 key metadata elements that would make the greatest impact on members' content making it useful to researchers and the entire scholarly community. These key elements add context and richness, and help to open up content to easier discovery and wider and more varied use. The 10 key elements Crossref chose to display In Participation Reports are: References, Open references, ORCID IDs, Funder Registry IDs, Funding award numbers, Crossmark-enabled, Text-mining URLs, License URLs, Similarity Check URLs, and Abstracts [3]. More detail on each is provided in the next section.

Why Are the 10 Key Elements in Participation Reports Important?

References

References are a big part of the story of a piece of content, highlighting its provenance and where it sits in the scholarly map. References give researchers and other users of Crossref metadata a vital data point through which to find content, which in turn increases the chances of it being read and used. They also enable members to use Crossref's Cited-by service, which means they can query for publications that cite a work, as well as showing citation counts and lists for articles (Fig. 2).

Open References

Open References displays the percentage of registered references that are set to be openly available. If a member has set their references to 'open' (and they are encouraged to do so), they're available to all users of all Crossref APIs and services. If not, fewer people can see and use them. Most members' citations are set to open but Participation Reports help easily check this and if the percentage is 0% then they are not set to open. Members registering references can make their references open by emailing Crossref's support team, and there is no charge to do so.

ORCID iDs

These persistent identifiers enable users to precisely identify a researcher's work—even when that researcher shares a name with someone else, or if they change their name. Governments, funding agencies, and institutions are increasingly seeking to account for their research investments. They need to know precisely what research outputs are being produced by the researchers that they fund or employ. ORCID iDs allow this reporting to be done automatically and accurately which is why Crossref encourages this. Adding ORCID iDs to Crossref metadata also enables ORCID auto-update, meaning that a researcher can be notified when a work connected with their ORCID iD is published, and they can choose to auto-

matically add that work and any future works to their ORCID profile, saving them time.

Funder Registry IDs

Funder Registry IDs identify organizations that funded the research. Publishers extracting these funding acknowledgements from content or collecting them via submission systems and adding them to Crossref metadata allows funding organizations to better track the published results of their grants, and allows publishers to analyze the sources of funding for their authors and ensure compliance with funder mandates.

Funding award or grant numbers

These are numbers assigned by the funding organization to identify the specific piece of funding (the award or grant). If funding award numbers are included then funding organizations are able to better track the published results of their grants and research institutions are able to track the published outputs of their employees.

Crossmark-enabled

The Crossmark service gives quick and easy access to the current status of a content item. With one click, a reader can see if the content has been updated, corrected, or retracted and

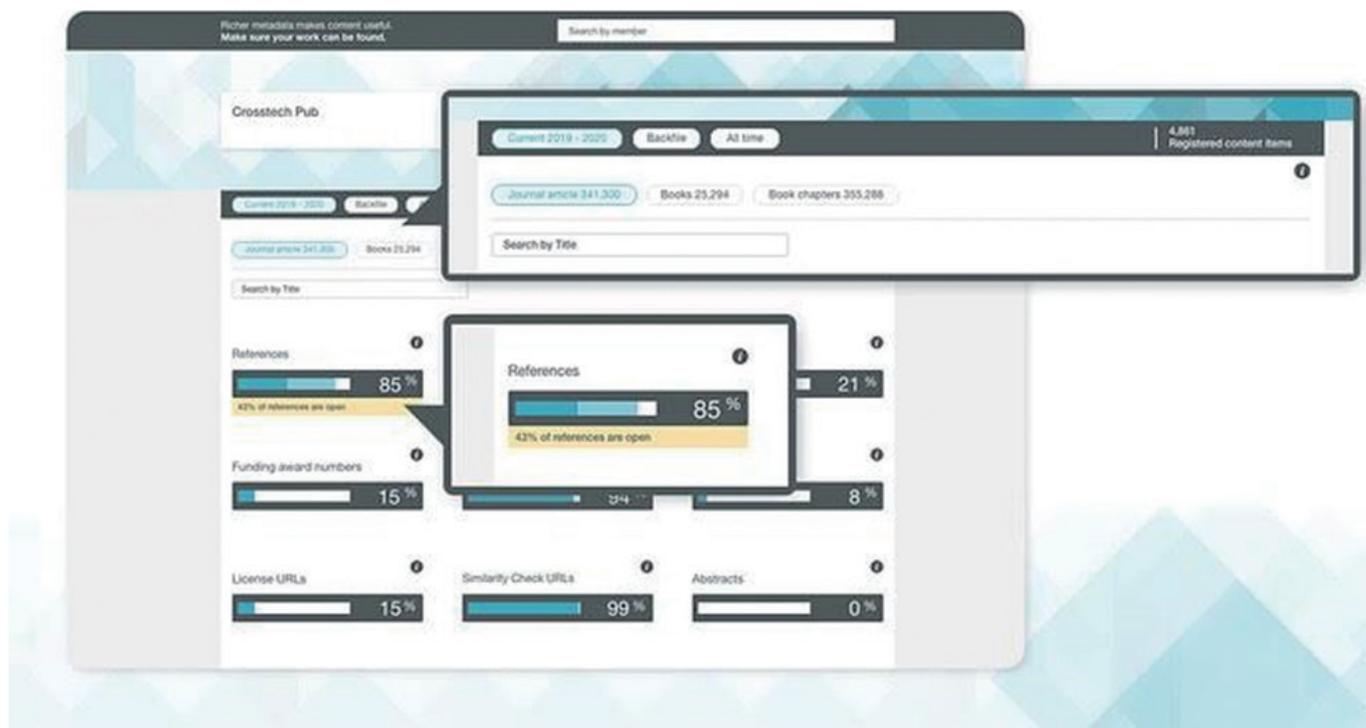


Fig. 2. Screenshot to show the percentage of content items that include reference lists in their metadata.

can access extra metadata provided by the publisher. It allows publishers to reassure readers that the publication keeps content up-to-date, and showcases any additional metadata the journal wants readers to view while reading the content (for example license and funding information or information on the peer review process).

Text-mining URLs

Researchers are increasingly interested in carrying out text and data mining of scholarly content, which is the automatic analysis and extraction of information from large numbers of documents. Text-mining URLs are links to the full text in the metadata (rather than just the landing page) to help researchers easily locate content for this purpose. Including full text URLs makes it easier for researchers to mine content, which increases discoverability and potential uses of the research.

License URLs

Members can include a link to their use and reuse conditions: whether this is their own proprietary license, or an open license such as Creative Commons. Including license URLs (or access indicators) in metadata is very helpful in letting readers know how they can access and use the content.

Similarity Check URLs

The Similarity Check service helps editors to prevent scholarly and professional plagiarism by providing editorial teams with access to Turnitin's powerful text comparison tool, and a comprehensive database of scholarly and other content to check documents against. Similarity Check URLs are full text URLs that enable iThenticate to index members' content into this database. Including Similarity Check URLs gives Crossref members access to the Similarity Check service, and also ensures that their content is included in these checks.

Abstracts

Including abstracts in the publication metadata gives more information to the user about a piece of content, making it more discoverable. Readers are more likely to navigate to an article if they can read an abstract because it gives further insight into the content of the work. Last year the I4OA (Initiative for Open Abstracts) was launched which encourages publishers to share their abstracts as part of their metadata in Crossref [4].

How to Use Participation Reports?

The reports are easy to use. Anyone can simply start by navigating to <https://www.crossref.org/members/rep/> then type in the member name into the search box and that will take them

to the report for that member [5]. The report dashboard page shows a variety of information including the total registered DOIs, content types, current or backfile content, and most importantly the 10 key metadata elements that are explained above. Next to the elements there are percentages that indicate what percentage of the DOIs include the particular metadata element. It's possible to filter by content such as journal articles, book chapters, datasets, and preprints, depending on what content types the member has registered. It's also possible to compare current content (past two calendar years and year-to-date) to back file content (older than that). And within the journal articles view, it's possible to drill down to view the metadata completeness for each individual journal. Crossref hears that editorial boards are keen to see that aspect!

Participation reports are free and open to everyone and don't require a login. Crossref recently agreed to adopt the POSI (Principles of Open Scholarly Infrastructure) [6], which offer a set of guidelines by which open scholarly infrastructure organizations like Crossref can be run and sustained [7]. And sharing metadata openly and investing in open infrastructure is one of the most important commitments that Crossref is trying to stand by.

How to Improve Coverage in Participation Reports?

Members should check their Participation Report and share it with their production teams or vendors to see what exactly is currently being registered with Crossref and what can still be added. Editors can encourage researchers to get and submit ORCIDs as part of the manuscript submission process. Funding data can be added from the acknowledgements sections. Organizations planning to join Crossref in future should make a plan to send as much metadata as possible to Crossref, focusing on the key elements listed in the reports.

Future Plans

Based on feedback from the community, Participation reports will see some updates in 2021. Some of these updates include (1) the member search bar will be incorporated on the dashboard page so that users will no longer need to go back to the Participation Reports homepage to find another member; (2) Crossref will make a few improvements to the member information displayed and will make the total registered content items display more accurately; (3) alternative or additional member names will be displayed; (4) refining how date ranges are changed (current content, backfile content, and all time); (5) simplifying filtering by content-type; and (6) Open References will be combined with References as a single key ele-

ment indicating percentage of open references all in one. Some members had found the current display confusing.

Crossref is also planning on adding additional key elements as they start to be collected via the metadata and become available in the REST API. ROR IDs (Research Organization Registry Identifiers) will hopefully be added next [8]. This persistent identifier connects research organizations to their outputs and makes it possible to see which researcher is working with which organization. In the future Crossref is also hoping to also add Grant identifiers which funders can now register with associated metadata. This will make it easier to include information about the use of facilities, equipment, salary awards and so on, and to show transparency into research funding and its outcomes.

Conclusion

The provision of rich metadata creates value for the research community. However, it's not always easy to see what important metadata Crossref members are registering for their publications. Crossref's Participation Reports provide an easy way to see who is registering what key metadata elements in Crossref. They can help members, authors, and editors figure out what important metadata elements are already registered and what's still missing. These elements make content more discoverable and useful, as they are used by researchers and the tools they use to help drive research further. Registering as much metadata as possible helps to make important research discoveries and connections that benefit the research community and the wider world.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Funding

The authors received no financial support for this article.

References

1. Crossref. Participation Reports [Internet]. Oxford: Crossref; 2020 [cited 2021 Jun 23]. Available from: <https://www.crossref.org/documentation/reports/participation-reports/>
2. Tolwinska A, Meddings K. 3,2,1... it's 'lift-off' for Participation Reports [Internet]. Oxford: Crossref; 2018 [cited 2021 Jun 23]. Available from: <https://www.crossref.org/blog/321-its-lift-off-for-participation-reports/>
3. Meddings K, Tolwinska A. How good is your metadata? [Internet]. Oxford: Crossref; 2018 [cited 2021 Jun 23]. Available from: <https://www.crossref.org/blog/how-good-is-your-metadata/>
4. I4OA. Initiative for Open Abstracts [Internet]. I4OA; 2021 [cited 2021 Jun 23]. Available from: <https://i4oa.org/>
5. Crossref. Participation Reports [Internet]. Oxford: Crossref; 2018 [cited 2021 Jun 23]. Available from: <https://www.crossref.org/members/prepare/>
6. Bilder G, Lin J, Neylon C. The Principles of Open Scholarly Infrastructure [Internet]. POSI; 2020 [cited 2021 Jun 23]. Available from: <https://doi.org/10.24343/C34W2H>
7. Bilder G. Crossref's board votes to adopt the Principles of Open Scholarly Infrastructure [Internet]. Oxford: Crossref; 2020 [cited 2021 Jun 23]. Available from: <https://www.crossref.org/blog/crossrefs-board-votes-to-adopt-the-principles-of-open-scholarly-infrastructure/>
8. ROR. Welcome to the Research Organization Registry community [Internet]. ROR [cited 2021 Jun 23]. Available from: <https://ror.org/>