



## Solutions for identification problems: a look at the Research Organization Registry

Rachael Lammey

Crossref, Oxford, UK

### Introduction

As DOIs and their associated metadata have been widely adopted in the research ecosystem to identify research outputs, there are still areas where identification of the full context or nature of these outputs is tricky, manual and therefore prone to mistakes. The first problem is one of identifying which organizations are affiliated with which research outputs. In other words, how can a research institution, like a university, easily see the publications and other outputs (such as data) that their researchers have produced? The second is identifying which grants resulted in which research outputs—how can a funder easily see the data, articles, preprints and more, connected to a specific project they funded? At the moment, reporting on both these areas is done manually, and relies on the researcher taking the time to let their institutions and funders know each time they publish an article. That time could be better spent doing research. There are also benefits for other stakeholders, including journals and publishers that having identifiers in place for organizations and grants will help accomplish.

### Introducing the Research Organizations Registry

Crossref is one of the four founding organizations behind the Research Organization Registry (ROR), a community-led project to develop an open, sustainable, usable, and unique identifier for every research organization in the world. The other organizations in the project team are California Digital Library, DataCite, and Digital Science. There is also a wider steering group with representatives from the Coalition for Networked Information (CNI), Japan Science and Technology Agency, the Association of Research Libraries and the Academy of Science of South Africa. The registry launched in January 2019 and began assigning ROR IDs to more than 91,000 research organizations using seed data from Digital Science's GRID database [1]. Since that time, the registry has expanded to cover over 97,000 entries. These can be found using the ROR registry search interface [2], or the ROR API [3].

Unique and persistent identifiers for the organizations affiliated with published research outputs minimize the amount of effort needed to disambiguate free-text affiliation fields or match outputs to affiliations. Adoption of ROR in scholarly infrastructure will enable more efficient discovery and tracking of publications across institutions and funding bodies, and will

**Received:** January 17, 2020

**Accepted:** January 19, 2020

**Correspondence to** Rachael Lammey  
rlammey@crossref.org

#### ORCID

Rachael Lammey

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

support existing workflows. Since there is such a wide range of use cases for ROR, it was felt that a collaborative approach between a number of organizations (all with their own use cases) would best serve the community.

### ROR in More Detail

The simplest way to show ROR is by looking at an entry in the ROR search interface (Fig. 1). In it, you will see the ROR ID for Yonsei University: <https://ror.org/01wjejq96>. The ROR ID is expressed as a URL that resolves to the organization's record. It is a unique and opaque character string: leading 0 followed by 6 characters and a 2-digit checksum. The record also shows crosswalks with other identifiers for the organization: GRID, International Standard Name Identifier, Crossref Funder Registry, and Wikidata. Other information shown includes the organization name in Korean (ROR supports multiple languages and character sets), the website, the location, and its primary purpose. All ROR IDs and metadata are provided under the Creative Commons CC0 1.0 Universal Public Domain Dedication [4], so they can be widely disseminated and used.

If there are issues with the entry in ROR, then organizations can contact [info@ror.org](mailto:info@ror.org) to request changes. Changes in ROR are currently implemented and passed to GRID, and GRID will also pass changes in GRID back to ROR.

### ROR in Use

ROR focuses on the “affiliation use case”—identifying which

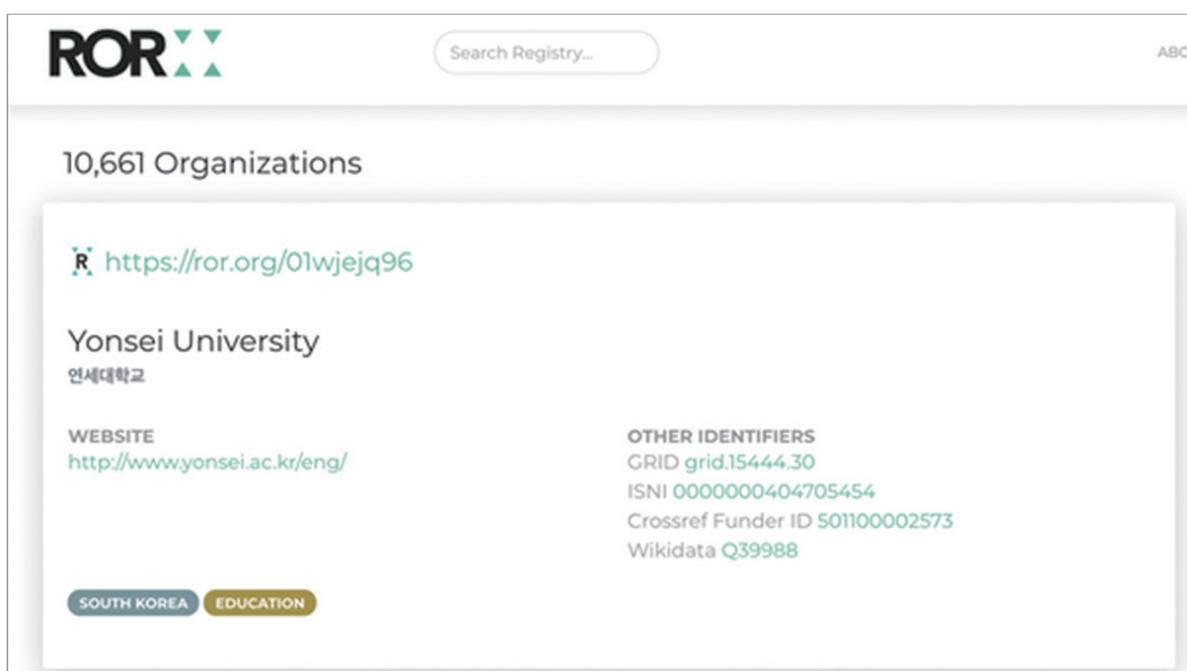


Fig. 1. The record for Yonsei University in the ROR (Research Organization Registry) search interface (<https://ror.org/search?query=yonsei+university>).

```

<journal_article publication_type="full_text">
  <titles>
    <title>
      Development of Ground Organic Monolith Particles as Packing Material in High Performance Liquid Chromatography
    </title>
  </titles>
  <contributors>
    <person_name contributor_role="author" sequence="first">
      <given_name>Yune Sung</given_name>
      <surname>Kim</surname>
      <affiliation>
        Department of ChemistryInha University Incheon 22212 South Korea
      </affiliation>
    </person_name>
  </contributors>
</journal_article>

```

Fig. 2. Author affiliation information in the current Crossref metadata schema.

organizations are associated with which research outputs. In this case, “affiliation” describes any formal relationship between a researcher and an organization associated with researchers, including but not limited to their employer, educator, funder, or scholarly society. In the case of journal publishing, editors will know that it can be difficult to clearly tell which institution an author comes from. Sometimes they will call the same institution different names, or an institution may change names or merge with another institution. As humans, we can look at the information and try to make an informed decision about the author or reviewer affiliation, but for the information to be available consistently, and at scale, the information needs to be readable in metadata too.

At Crossref, for example, information on the author’s institution is often missing from the article metadata. If the publisher does include it, they can only do so in a text string (Fig. 2). The information is clear, but entering the full text is open to error, both when the author inputs it and when the article is being prepared for publication. This also makes the information hard to report on or analyse at scale, because there can be so many variations in how an affiliation is entered. Crossref will start to accept ROR IDs in the metadata they collect later in 2020, with current proposals [5] suggesting ROR should be included as shown in Fig. 3.

Once Crossref starts to accept ROR IDs in publisher metadata, they will be disseminated via Crossref’s APIs and can be used by publishers to do analysis of where their authors are coming from, and by the institutions themselves to get a more comprehensive picture of what their researchers are publishing and where they are publishing it.

Some organizations have already integrated ROR IDs into their workflows. One example is Dryad, an international repository of research data. In Dryad, when an author starts to enter their affiliation (shown in Fig. 4), the lookup tool searches for a matching name in ROR and shows the author a drop-down list of possible matches to choose from. This will work regardless of whether the author starts entering a known abbreviation or the full name of the organization.

Dryad registers their DOIs with DataCite, and includes the ROR IDs with their DOI deposits so that they are searchable via DataCite’s search interfaces. The collection of this information by Dryad will ensure that research organizations will be able to find what data their researchers are making available via that repository. Other implementations of ROR are in progress which will help add ROR IDs in other research workflows, for example, Altum who run a system used by many funders to manage their grant information.

```
<affiliations>
  <institution>
    <institution_name>University of Somewhere
Awesome</institution_name>
    <institution_acronym>USA</institution_acronym>
    <institution_id
institution_id_type="ror">https://ror.org/02twcfxyz</institution_id>
    <institution_place country="ca">Winnepeg</institution_place>
    <institution_department>Feline Research</institution_department>
  </institution>
</affiliations>
```

Fig. 3. Suggested format for the collection of Research Organization Registry IDs in the Crossref Metadata Schema.

The screenshot shows a web form titled "Dataset: Basic Information". It includes fields for "Title" (My Dataset), "Author(s)" (First Name: Maria, Last Name: Gould), and "Institutional Affiliation". The "Institutional Affiliation" dropdown menu is open, showing a list of options: "CDL", "California Digital Library", "CdLS Foundation UK & Ireland", "Conseil de Developpement du Loisir Scientifique", and "Cornelia de Lange Syndrome Foundation". The "CDL" option is highlighted in orange. An "Author Email" field contains "email@example.com". A URL "https://sandbox.orcid.org/0000-0003-0323-450X" is visible below the author information, along with a "+ Add Author" button.

Fig. 4. Research Organization Registry integration in Dryad.

## Introducing Global Grant Identifiers

2019 saw the funder Wellcome register the first Grant IDs with Crossref. These persistent identifiers for research grants (rather than just the research funder) will help Wellcome and other funders easily find out what their researchers are publishing, where they are publishing and how - for example, many funders now require researchers to share their research data and to publish open access. Once Grant IDs are provided by the funder and passed into the publication metadata, the funder will then be able to search for a Grant ID and then see what has been associated with that grant. This will help them understand connections between projects and collaborators, mean less duplication of effort in overlapping grants or repeated projects and identify pockets of expertise and emerging areas of activity.

Grant IDs will help publishers generate more accurate funding acknowledgements, make sure that content is being published in line with funder policies, help reviewers spot potential conflicts of interest, and provide another mechanism for publisher content to be discovered and used.

## How Grant IDs Are Implemented

Bilder and Hendricks [6] explains that a funder registering metadata and creating DOIs for grants would need to take the following steps: first, when a grant is submitted, the funder would assign their own internal identifier for tracking, e.g., 05-67-89. Most funders already use these internal numbers. Second, if the grant is accepted, the funder would 1) generate a global public identifier for the grant based on the DOI. For example, assuming their prefix was 10.4440, then the global public identifier could be <https://doi.org/10.4440/00-00-05-67-89>; 2) create a “landing page” to which the identifier will resolve. The landing page would display metadata describing the grant, as well as a link to the grant itself; and 3) register the generated DOI and metadata with their registration agency (e.g., Crossref or DataCite).

As with DOIs, if the metadata related to the grant changes, the funder would then update the information with Crossref. The funder would also promote the use of the DOI as the global identifier for the grant, and ask people to use it when referring or citing the grant. For example, authors would enter this information into submission systems in the field where they currently enter the award or grant number. The information would then be placed in the corresponding field in the Crossref metadata which is made available once the paper is published.

Crossref is working on this initiative with a wide group of funders including those from its Funder Advisory Group [7]. They have advised on the project from the start and will be

among the first funders to register Grant IDs with Crossref. Once a critical mass of these identifiers and metadata are registered, they will start to populate through publishing workflows. Crossref is talking to various stakeholders including manuscript submission systems, hosting platforms, publishers and metadata users to make sure they are briefed on how they can integrate Grant IDs into their tools and services.

## Conclusion

More stakeholders in the research process are getting involved in publishing workflows. They see an advantage in using these to try to fill in gaps in the map of the research landscape with new data points and better quality information to inform their policies and help them automate reporting to try to save their researchers time and effort. Persistent identifiers, such as ROR IDs and Grant IDs, if they are well-adopted and used, will help them with this. It is still relatively early in the lifespan of both initiatives, but they serve the broad needs of the research community and will expect their takeup to grow quickly in the next number of years.

## Conflict of Interest

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

## References

1. GRID [Internet]. [place unknown]: Digital Science & Research Solutions [cited 2020 Jan 15]. Available from: <https://grid.ac/>
2. Research Organization Registry. ROR search [Internet]. [place unknown]: Research Organization Registry [cited 2020 Jan 15]. Available from: <https://ror.org/search>
3. Research Organization Registry. ROR API [Internet]. [place unknown]: Research Organization Registry [cited 2020 Jan 15]. Available from: <https://api.ror.org/organizations>
4. Creative Commons. CC0 1.0 universal (CC0 1.0) public domain dedication [Internet]. Mountain View, CA: Creative Commons [cited 2020 Jan 15]. Available from: <https://creativecommons.org/publicdomain/zero/1.0/>
5. Feeney P. Proposed Schema Changes: have your say [Internet]. Oxford: Crossref; 2019 [cited 2020 Jan 15]. Available from: <https://www.crossref.org/blog/proposed-schema-changes-have-your-say/>
6. Bilder G, Hendricks G. Global Persistent Identifiers for grants, awards, and facilities [Internet]. Oxford: Crossref; 2017 [cited 2020 Jan 15]. Available from: <https://www.>

[crossref.org/blog/global-persistent-identifiers-for-grants-awards-and-facilities/](https://crossref.org/blog/global-persistent-identifiers-for-grants-awards-and-facilities/)

7. Crossref. Funder advisory group [Internet]. Oxford:

Crossref; 2019 [cited 2020 Jan 15]. Available from: <https://www.crossref.org/working-groups/funders/>