Capturing use of research facilities with PIDs

PIDapalooza,
Girona, Spain,
23 January 2018

Erin Arndt | Wiley | https://orcid.org/0000-0002-9012-1425

Laure Haak | ORCID | https://orcid.org/0000-0001-5109-3700

Crystal Schrof | Oak Ridge National Lab | https://orcid.org/0000-0002-8462-4579

Susan White-DePace | Argonne National Lab/Society for Science at User Research Facilities |
https://orcid.org/0000-0002-5031-4935

Imagine a world where:

Researchers are easily and authoritatively connected with funding, affiliations, equipment, research facilities, datasets, samples, and publications!

This is **NOT** Science Fiction!

Issue

- Government-funded laboratories award researchers use of facilities and equipment based on a competitive proposal process

- Government agencies support collections of samples used by researchers

- **Making accurate assessment of the scientific impact of these public investments difficult, if not impossible**

Background

- Oak Ridge National Laboratory's (ORNL) neutron sources began collecting researchers' ORCiD in early 2016 to:
  - Calculate the return on investment for outreach activities
  - Acknowledge award of user facility resource and facility service
  - Connect user facility to researcher's publication through ORCiD

Background

- The Society for Science at User Facilities (SSURF) Annual Meeting focuses on PIDs and data sharing in June 2016 and recommends a pilot training reciprocity program using ORCiDs

- ORNL, DOE's Office of Science and Technical Information, and ORCiD discuss a project to create an identifier-enabled workflow for user facilities

- In January 2017, DOE, national laboratories, SSURF, CHORUS, publishers, and ORCiD established a working group to define terms, user stories, and information flows that *leverage open identifier infrastructure*

- The WG published a Report in November 2017 outlining findings and recommendations

- A series of proof of concept pilot projects will launch in 2018

User Facilities and Publications WG: Findings and Opportunities

Authors: Laurel L Haak, Erin Arndt, Benjamin Brown, Mark D. Doyle, Mariam Elsayed, Patricia Garvey, Terry Law, Kathleen Nasta, Robert Peters, Howard Ratner, Crystal Schrof, Steven Watson, Susan White-DePace

Abstract: User facilities are specialized government-sponsored research infrastructure available for external use to advance scientific or technical knowledge. Researchers compete for access to these facilities and specialized equipment, but do not consistently acknowledge the contributions made by facilities when they publish the results of their work. Collection of these research outputs is necessary to enable an accurate assessment of the scientific impact of these public investments. Recently several user facilities have begun to explore innovative approaches to these challenges, including adoption of ORCID.

Working Group

- Erin Arndt, Wiley (US)

- Ben Brown, U.S. Department of Energy (DOE) Office of Science (US)

- Josh Brown, ORCID (UK)

- Mark Doyle, American Physical Society (US)

- Mariam Elsayed, DOE Office of Science (US)

- Laure Haak, ORCID, Chair (US)

- Stephanie Harley, ORCID, Project Manager (US)

- Ted Hatcher, American Chemical Society (US)

- Matthew Hawkins, National Science Foundation (US)

- Terry Law, Pacific Northwest National Lab (US)

- Mark Martin, DOE OSTI (US)

- Kathy Nasta, Brookhaven National Lab (US)

- Dan O'Brien, American Chemical Society (US)

- Rob Peters, ORCID (US)

- Howard Ratner, CHORUS (US)

- Crystal Schrof, Oak Ridge National Lab (US)

- Steve Watson, Elsevier (Netherlands)

- Susan White-DePace, Argonne National Lab/Society for Science at User Research Facilities (US)

Findings

- Even when required by policy, researchers often do not report the facilities or specialized equipment used to develop papers, datasets, or other research results

- Publishers do not collect information on user facilities, instead focusing on funding sources for the published work that are related to journal article access policies

- Differing definitions of terms such as author, user, and funding among stakeholders present challenges to effective reporting

- Stakeholders see value in using APIs and persistent identifiers including ORCiD and DOIs to enable reporting processes

Working Group began by establishing workflows

User stories describing interactions between users and scientific facilities, from the time a researcher requests time at a facility, to the time they publish a paper based on work done at that facility were developed

- STEP 1: Request use of facility
- STEP 2: Approve facility use
- STEP 3: Researcher performs work at facility
- STEP 4: Submit article describing research
- STEP 5: Article is published
- STEP 6: Reporting

Facilities Workflows

Step 1. Proposal Submission

Researcher has an idea for a project, and determines that specialized resources available at a facility are required to carry out the project. Researcher submits a proposal to the facility. The facility collects the proposal and assigns it to reviewers. The facility:

- Collects ORCiDs and leverages the ORCiD record to obtain researcher host institution and organization ID from the proposal submitter and co-proposers

- Collects ORCiD and leverages the ORCiD record to obtain home organization name and organization ID from the reviewer(s) of proposals

- Requests permission from the proposal submitter and co-proposers to update their ORCiD record should the proposal be awarded

- Assigns a unique identifier to the proposal

Step 2-3. Proposal Accepted and Facility Use

Based on the review of the scientific merit of the proposal, the facility decides to approve the project. Researchers are granted time at the facility and use of specialized equipment as detailed in their proposal. The facility:

- Obtains an external Proposal/Project PID for the facility access award

- Posts public information about the award in a public database

- Uses the ORCiD permission token to post award metadata (Proposal/Project Identifier, facility name, and facility organization ID)

- Considers collecting ORCiDs, researcher host organization IDs, and IDs in and for data management plans

Pilot Project 1: Integrate identifiers into user facility registration and proposal processes

- Collect ORCiDs

- Make the award citable

- Facilitate information sharing

- Reduce reporting burden

It works! ORNL has implemented several of the workflows outlined in Steps 1 and 2

Jingxuan Ding

ORCID ID

ⓘ https://orcid.org/0000-0003-3126-7165

🖨 Print view ❓

⌄ **Funding (1)**   ⏬ Sort

Neutron Beam Award at Spallation Neutron Source (SNS)

Oak Ridge National Laboratory (Oak Ridge, TN, United States)

2017-01 to 2017-06 | Award

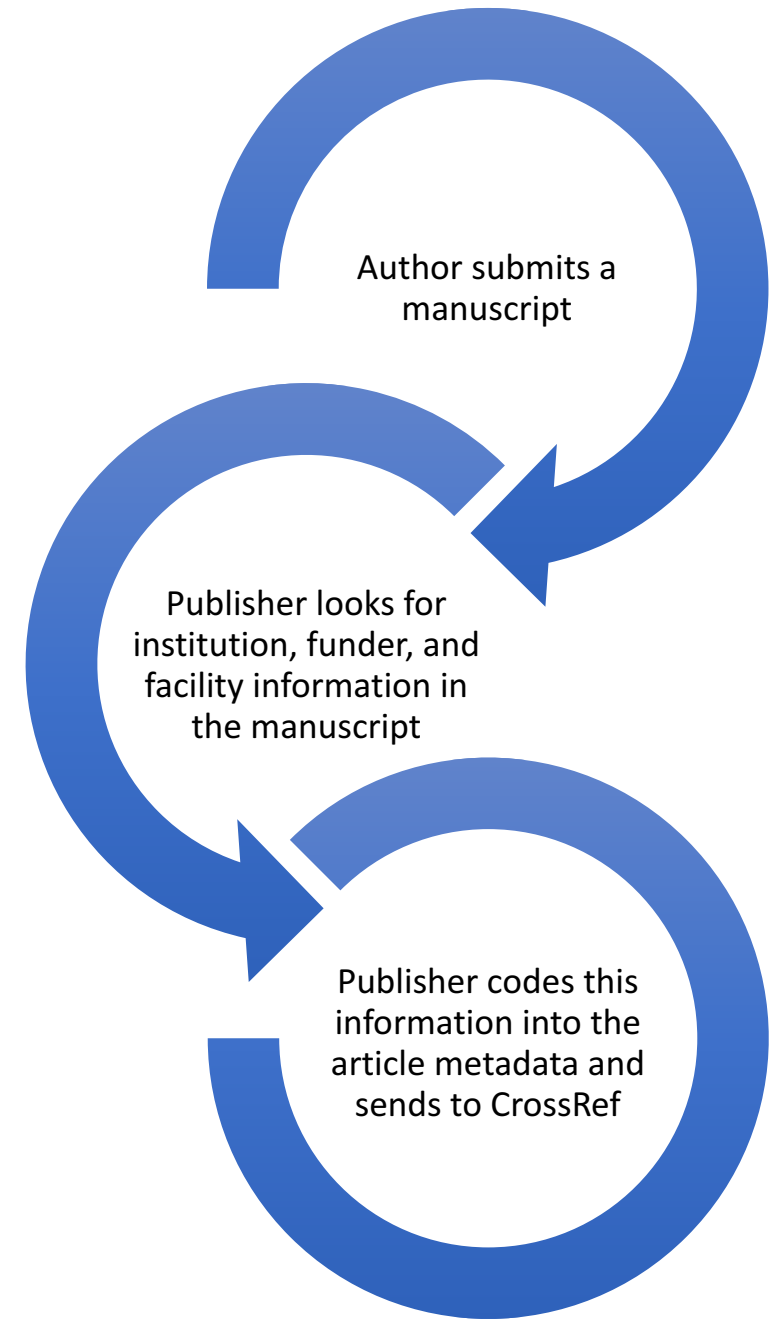Award: 2017-A: CNCS

Alternate URL: neutrons.ornl.gov

Source: Oak Ridge National Laboratory   ☑ Preferred source

Interested in piloting a project at your organization?
Contact community@orcid.org for more information.

Publisher Workflows

Current State

- Publishers receive information from authors about facilities in the manuscript, typically in the Acknowledgements section, with no PIDs or standards applied.
  - This information may be parsed either at authoring, submission, a revision stage, or directly before publication
- Some publishers additionally ask authors to identify their institutions and funders using tools within editorial systems where a PID can be automatically applied
  - There is currently no solution of this type for facility information

Author submits a manuscript

Publisher looks for institution, funder, and facility information in the manuscript

Publisher codes this information into the article metadata and sends to CrossRef

Pilot Project 2: Integrate award and facility IDs into the manuscript publication process.

- Collect ORCiDs at time of manuscript submission or acceptance.
- Couple collection of ORCiDs with other data from the author's ORCiD record
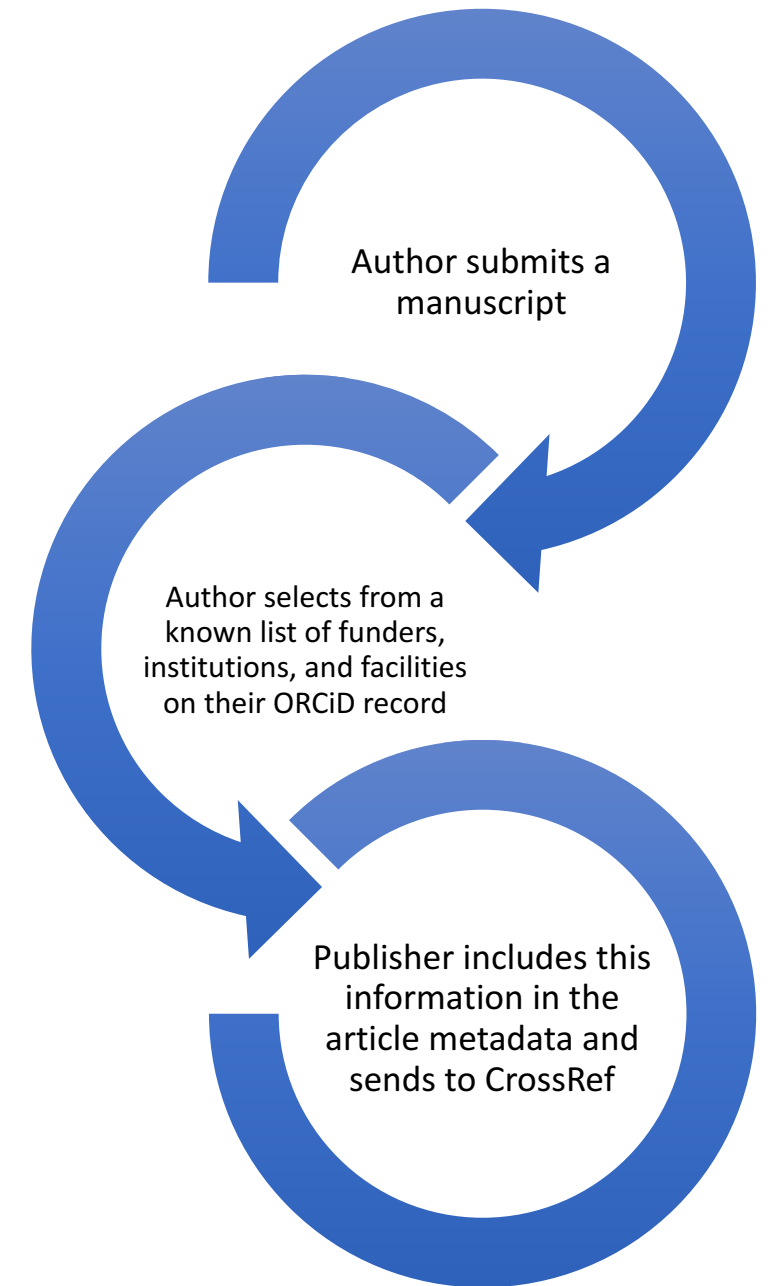- Amend JATS standard to enable collection of facility use information.

The Plan

- **Path 1:** Work with editorial office systems on a solution that ingests facility information directly from a user's ORCiD profile

- **Path 2:** Build a stand-alone widget that pulls facility information directly from a user's ORCiD profile into non-editorial office system workflows, such as proofing or license signing, where a publisher can choose when they'd like to receive this data

- **Note1:** Either plan would also apply to ingesting funder and institutional data

- **Note 2:** An update to the Journal Article Tag Suite (JATS) is required to support facility information

Facilities and JATS

```xml
<research-support-group>
    <award-group award-type="approved-proposal">
        <research-support-source support-type="user-facility">
            <institution-wrap>
                <institution>Spallation Neutron Source</institution>
            </institution-wrap>
            <institution-wrap>
                <institution>Oak Ridge National Laboratory</institution>
                <institution-id institution-id-type="doi" vocab="open-funder-registry" vocab-
                identifier="10.13039/open_funder_registry">10.13039/100006225</institution-id>
            </institution-wrap>
                <research-support-description>Beam time and computing resources</research-support-description>
        </research-support-source>
        <principal-award-recipient>
            <name>Dr. Albert Einstein</name>
            <contrib-id contrib-id-type="orcid" authenticated="true">http://orcid.org/0000-0000-0000-0000</contrib-id>
        </principal-award-recipient>
        <award-id>SPS 12345</award-id>
    </award-group>
    <research-support-source support-type="research-materials">
        <research-resource-wrap>
            <research-resource>Slc9a4 (C05) tm1b Mus musculus</research-resource>
            <research-resource-id resource-id-type="rrid" vocab="Research Resource
            Identifier">RRID:IMSR_HAR:5669</resource-resource-id>
        </research-resource-wrap>
    </research-support-source>
</research-support-group>
```

Future State

- Authors and co-authors are presented with a list of all facility, funder, and institutional information (incl. PIDs) associated by those parties to their ORCiD record.

- They can select which information to apply to a given piece of research for publication, but additionally report other information not found in the ORCiD record
  - If a dataset was created and published to a repository, the DOI for the dataset can also be pulled into the publication metadata.

- The publisher will include this information (and the underlying PID) in the publication metadata, and feed it back to CrossRef (and thus ORCiD) so that facilities, funders, and institutions are aware that they were properly acknowledged.

Author submits a manuscript

Author selects from a known list of funders, institutions, and facilities on their ORCiD record

Publisher includes this information in the article metadata and sends to CrossRef

Interested? Join us!

- Get in touch: community@orcid.org
- Consider participating in a pilot project