

# Key Issue

## Collective Action for the Open Researcher & Contributor ID (ORCID)



**MARTIN FENNER**

Clinical Fellow  
Hannover Medical School,  
Germany



**CONSOL GARCIA  
GÓMEZ**

Librarian  
Universitat Politècnica de  
Catalunya, UPC-BCN Tech,  
Barcelona, Spain



**GUDMUNDUR A  
THORISSON**

Research Associate  
University of Leicester, UK  
and University of Iceland,  
Reykjavik

*The authors are members of the ORCID Outreach Working Group*

The Open Researcher & Contributor ID (ORCID) initiative was started in November 2009 to solve the author name ambiguity problem in scholarly communication by establishing a global, open registry of unique identifiers for researchers<sup>1</sup>. ORCID differs from other author identifier services in that ORCID will not be limited by discipline or geographic region, and more than 260 organizations worldwide are already participating in this international initiative. ORCID is guided by a set of principles that stress the openness of the organization<sup>2</sup>.

The ORCID software is based on Researcher ID code licensed from Thomson Reuters and will be released as open source software when the first round of development is completed. The ORCID query and deposit application programming interfaces (APIs) will be available in November 2011. These early APIs will initially only return mock data, but their main purpose is to allow developers to start integrating ORCID into their systems in anticipation of the launch of the public service. ORCID will launch a public registration

service for individual researchers in March 2012, at which time ORCID will also start to offer paid membership to organizations (registration will always be free for individual researchers). Later in 2012, ORCID will start working on the next major version of the service which will facilitate sophisticated disambiguation of claims (assertions about authors and authorship) imported from multiple sources, e.g. publishers and funding organizations.

### A collective action problem

The success of ORCID depends on a critical number of enabling services and users. It is the perfect example for a collective action problem, described in detail by economist Mancur Olson in 1965<sup>3</sup>: without selective incentives for participation, collective action is unlikely to occur even with large groups of people with common interests. What this means for the forthcoming launch of the ORCID service is that ORCID has to focus on incentives for individual groups of stakeholders, and that the adoption will

happen in stages, adding value for a particular group at each stage. In this issue we describe four distinct stages of key importance in the anticipated adoption process of the ORCID service. The intention is to provide a better understanding of how and when ORCID could provide a valuable service to a particular group of stakeholders.

A unique author identifier is only really useful when used for claims about biographic and bibliographic information. These can be self-claims by the author himself, or external claims by other members of the scholarly ecosystem, including universities, funding organizations and publishers. A good number of authors are keen to join ORCID as soon as the service launches, but most will join ORCID when they submit a manuscript, data set or grant proposal, or when their institution starts using ORCID. It is helpful to distinguish between active and inactive authors, and between unpublished and already published scholarly content (Table 1). Whereas the effort to associate scholarly content with its authors increases from top to bottom, the added value decreases in parallel. It therefore makes sense for ORCID to first focus on works that can be claimed at the time of publication, and on claims already made by authors and external sources.

### Stage 1 – unpublished scholarly content

Using ORCID identifiers at the time of publication immediately creates value for publishers of scholarly papers and other scholarly content, including data sets and books. ORCID will streamline the manuscript submission and review process, and will decrease the costs for publishers and the work required for authors. ORCID will be of particular value for data publishers, as ORCID will help with adoption of the emerging concept of data citation. ORCID identifiers create similar value for funding organizations, who can use ORCID identifiers in grant applications and grant reporting. For these reasons, funding organizations and publishers of journal articles and other scholarly content will probably be among the first paying ORCID

members, and they will be crucial for the early adoption of ORCID identifiers by authors.

### Stage 2 – published and claimed scholarly content

ORCID is obviously not the first attempt to create scholarly author identifiers. Other services have for many years allowed claims about authors and their publications. ORCID is special, however, because the service will be not only much broader in scope, but also allow authors to link their ORCID profile to other author identifier services, and to import claims from these services. Individual authors will be very interested in this profile linking and import if the ORCID service provides unique features not currently available. The most requested (but by no means the only) feature is an enhanced author profile that can be used as template for institutional homepages and CVs. Another popular feature (only partially covered by other author identifier systems) is the inclusion of scholarly contributions other than journal articles and books in the ORCID profile, including but not limited to data sets, peer review and microattributions<sup>4</sup>.

The co-operation of the existing author identifier services is important in getting the ORCID service off to a good start. Not only do these services already contain large numbers of claims, but they are also often built around active researcher communities. ORCID is therefore working closely with most of these services from the outset, and the upcoming ORCID APIs will enable exchange of profile information between ORCID and other author identifier services. ORCID will allow both the import and export (via the deposit and query APIs, or via mass ingest) of profile information, enabling the exchange of information in both directions. However, wide-ranging exchange of profile information could potentially create large disambiguation problems. Until the ORCID disambiguation software is completed in 2013, profile exchange will therefore rely on a very close co-operation of all involved parties, and is best done by directly involving the authors themselves.

<b>Stage 1</b>	active author	claimed	unpublished content
<b>Stage 2</b>	active author	claimed	published content
<b>Stage 3</b>	active author	unclaimed	published content
<b>Stage 4</b>	inactive author	unclaimed	published content

Table 1. Anticipated adoption stages of the ORCID service

### ORCID reaches critical number of claims

When enough scholarly works have been published with an ORCID identifier, and when sufficient claimed records have been imported into the ORCID system, the service will begin to provide value to universities and research institutions. Organizations can start to automatically and accurately track scholarly content produced by their researchers. This will not only greatly lower the effort and costs required to perform evaluation of their researchers, but will also make it easier for organizations to conform to open access mandates and other reporting requirements. Some institutions already have established research information systems and might be interested in early ORCID adoption, but many others will probably wait until enough scholarly content is automatically pushed into the ORCID system before they start using the service. Most universities and research institutions will therefore probably not become paying ORCID members until at least 2013.

### Stage 3 - published but unclaimed scholarly content

Compared to newly created scholarly content, it is more difficult to make claims about the authors of already published scholarly content. Fortunately, for scientific research purposes this is also less important, especially for content that is more than a few years old. Nonetheless, automated author name disambiguation – a process that is known to be at best 90-95% accurate – can greatly facilitate this process, especially if results can then be verified by the authors (if still active, see below) or other parties acting on their behalf. This stage is of particular interest to bibliographic databases, as ORCID can help them to verify the association of authors with content in their own databases. The first major update of the ORCID software scheduled for 2013 will focus on disambiguation.

### Stage 4 - inactive authors

Uniquely identifying scholarly content produced by authors that are no longer active or already deceased is even less important and also more

complicated. This stage will depend largely on automated disambiguation. For these and other reasons, ORCID will initially focus on active researchers, and creation of ORCID identifiers and disambiguated author profiles for inactive authors will be tackled at a later stage.

### You can make this happen!

The outline above is not meant as a timeline, as these stages will partly take place in parallel and are difficult to predict in detail. But almost two years after the first ORCID announcement in December 2009 the expectations are high, and it is important to stress that there will be no 'day one' when the ORCID service is simply switched on. Launching a new global author identifier service is a unique challenge – ORCID wants all stakeholders to understand that it requires a significant and co-ordinated effort in the next 12 months to make this work.

### References

1. Fenner, M, ORCID: Unique Identifiers for Authors and Contributors, *Information Standards Quarterly*, 2011, 23(3). Available at <http://dx.doi.org/10.3789/isqv23n3.2011.03> (accessed 3 October 2011).
2. ORCID Principles: <http://www.orcid.org/principles> (accessed 19 September 2011).
3. Olson, M, *The logic of collective action*. Harvard University Press, 1965.
4. Giardine, B, et al, Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach, *Nature Genetics*, 2011, 43(4), 295–301. Available at <http://dx.doi.org/10.1038/ng.785> (accessed 3 October 2011).

Key Issue © Martin Fenner, Consol Garcia Gómez and Gudmundur A Thorisson

#### Contact:

Martin Fenner

E-mail: [fenner.martin@mh-hannover.de](mailto:fenner.martin@mh-hannover.de)