

Identifying Researchers: an Initial Requirements Statement

An often-stated requirement of researchers, research managers and the research library and repository community is for an overarching framework that would enable them to unambiguously identify researchers and their research contributions.

In order to establish an initial starting point for defining these requirements, the Workshop organisers have commissioned a set of requirements for public research identifiers for discussion.¹

Definitions

According to Kim Cameron, digital identity is “a set of claims made by one digital subject about itself or another digital subject.”²

A researcher identity is a special case of digital identity that focuses on the set of public claims that can be made about a researcher, facilitated by a researcher identifier.

For example, the research identity associated with identifier *R* might be:

	is	has
Identifier <i>R</i>	Author of Book <i>W</i>	name <i>A</i>
	Chief Investigator of ARC Grant <i>X</i>	GAMS ³ ID <i>B</i>
	Employee of Institution <i>Y</i>	Institution ID
	Creator of Dataset <i>Z</i>	email address <i>C</i>

The researchers themselves will have made some of these claims, whilst colleagues, research administrators, repository managers and so on will have made others on their behalf.

Many of these claims will be stored in institutional information systems (e.g. human resource and finance systems); others will be stored in public directories and registers. It is also often the case that this information is not shared within or across institutions, or with third parties, even when security and privacy issues are accounted for.⁴

In an ideal world, researchers would be consistently and uniquely identified against all publications, created works, and funded research for which they are associated with.

As public identifiers, researcher identities have particular requirements:

Requirements for identity control and management

1. Researchers should have control over their identifier, as the use of this identifier shapes what is publicly knowable about them.
2. For a given research item or output created in a given context, researchers should be aware of (and have control over) what identifier should be assigned.
3. There should be general agreement about what sort of claims about a researchers

Discussion papers

identity can be made using a researchers own public identifier.

4. Researchers should not be able to assert that they are somebody else (or supply somebody else's unique identifier).
5. It should be difficult for a third party (such as a repository manager) to incorrectly attribute a work to the wrong person. If incorrect statements are made about a researchers identity, it should be clear who the researcher needs to contact to redress the error and a documented process in place to ensure it is corrected.
6. Research Identifiers should be global and not limited to gated communities.
7. Research Identifiers should be portable; that is, they should reside with the researcher for the duration of their research career.

Requirements for collaboration

8. Researcher identifiers should be easy to communicate by the researcher, and therefore used in many collaboration contexts

Requirements for provisioning

9. It should be easy to establish a collaborative space for researchers who may work in universities and research organizations all over the world.
10. Researchers should be registered in a global federated directory and should be able to be selected from it
11. In general, a users public research identifier should be used in determining what access a user has to digital resources within a system; however, it should not be used to determine whether or not a researcher should have access to the system in the first instance.
12. Public research outputs from a research collaboration should uniquely and unambiguously identify the collaborators

¹ Simon Porter, Information Manager (Research), eScholarship Research Centre, University of Melbourne, authored the requirements.

² "The Laws of Identity" [Kim Cameron, Identity and Access Architect, Microsoft Corporation, May 2005] <http://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf>

³ Grant Application Management System (GAMS):
http://www.arc.gov.au/applicants/gams/gams_default.htm

⁴ Australian Government e-Authentication Framework (AGAF) for Individuals, Discussion Paper. December 2005. Australian Government Information Management Office (AGIMO)
http://www.agimo.gov.au/_data/assets/pdf_file/0005/46796/AGAF_I_Discussion_Paper.pdf

Six Use Cases

Recognition of the data researcher

A Researcher wants to be able to retrieve usage statistics for research data that they have deposited in multiple repositories. To initiate the query they supply their researcher identifier, which is then used to identify their research data, and gather usage statistics

People Australia

Universities supply information to the National Library of Australia about publications produced by their researchers. Researchers are identified by their researcher identifier, which is compared with the Library's Name Authority records. This is aggregated along with other information held by the National Library on the individual to provide the public with a profile based on aggregated public resources.

Provisioning

A new collaboration is being established. Using a collaboration tool, one of the researchers sets up a project site. Using a global directory of Researchers, other researchers are selected and added to the project.

There is a mechanism for the other collaborators to gain access to the collaboration space based on validated assertions of what their identifier is.

Output from collaborations

As researcher data from a collaboration is created and placed into repositories, the unique public identifiers for the associated researchers are included as part of the metadata of research output.

Statements of identity by third parties

A librarian, entering a new publication into a repository (or elsewhere) is able to unambiguously and uniquely identify authors from their own and other Universities using a global directory of researchers.

As publications are submitted to journals, authors are uniquely identified according to their researcher identifier.

Mobility

A Researcher arrives in a new University. Upon supplying their researcher identifier, metadata about their publications is downloaded into the system for use in their public staff profile and the university's research register.