Project THOR: Persistent Identifiers in the service of open, interoperable research

Tom Demeranville ORCID EU/THOR, Josh Brown ORCID EU/THOR

Keywords
Persistent Identifiers, Interoperability

Summary

Project THOR works with persistent identifiers across discipline and infrastructure boundaries to increase interoperability and promote open science. THOR aims to improve links between article, data, funder, institution and researcher metadata throughout the research lifecycle.

This session will demonstrate the value of persistent identifiers, describe work done so far to increase their utility and outline future plans. It will also explain how to get involved and realise the benefits of persistent identifier use within your own organisation.

The THOR consortium includes ORCID, DataCite, EMBL/EBI, The British Library, CERN, PANGAEA, DRYAD, Elsevier Labs and PLOS. It is a 30 month project funded by the European Commission under the Horizon 2020 programme.

Extended Abstract

The THOR project (http://project-thor.eu) is a 30 month project funded by the European Commission under the Horizon 2020 programme. THOR aims to extend the integration of persistent identifiers (PIDs) into platforms, services and workflows. This will place PIDs at the fingertips of researchers and institutions, building them into the services that they already use, across disciplines, regions and systems. The aim is not to build new, standalone services, but to work with existing systems and the community to increase interoperability, the use of PIDs and to solve shared problems. By creating new, and improved, integrations of PIDs in the services that researchers and institutions actually use, we aim to ensure that PIDs are embedded in research outputs and activities from the very beginning, with no additional effort for researchers.

Persistent Identifiers (PIDs) uniquely identify entities within the research ecosystem and help define relations between contributors, research artefacts, and organizations. They can be used to link across disciplines and infrastructures to build clearer pictures of how research is generated. They not only provide connections between authors and articles, but also between datasets, funders and institutions, enabling better attribution and information on where research originated. PIDs are a vital part of e-infrastructure that enable open and reproducible science and enhance the exchange of research information through interoperable systems.

Information about researchers' activities is gathered in multiple systems outside researchers' home institutions. Manuscript submissions systems, grant funding applications, datacentres, citation indices, other institutional or disciplinary repositories and personal webpages are all vital sources of information. PIDs enable the discovery and collection of this information, and ensure that data can be compared, matched and combined with greater efficiency and accuracy. This is vital now, as institutions work to understand their entire research portfolio. PIDs can help make research reporting more efficient (borne out by the growing number of funders mandating ORCID iDs, for example) and can help to demonstrate compliance with policies for research data management and open access.

This importance is only likely to grow as PIDs become embedded in even more systems, and linked to new kinds of information (about software, data analysis or peer review for example) meaning that institutional systems that do not make full use of these tools as they emerge risk under-serving the institution and its researchers. THOR partners are not just working to increase the availability of PIDs to researchers and the systems that support them, they are building tool kits, resources and training programs to ensure that information and technology professionals across the sector are ready to make the best use of PIDs every day.

The THOR consortium includes ORCID, DataCite, EMBL/EBI, The British Library, CERN, PANGAEA, DRYAD, Elsevier Labs and PLOS. Alongside publishers, datacentres, research organisations and national libraries, there are two major PID infrastructure providers; ORCID and DataCite.

ORCID provides identifiers for people involved in generating research. These identifiers are persistent, actionable and provide links to research activities such as publications, datasets and funding via other identifier systems such as Digital Object Identifiers (DOIs) and Fundref. ORCID records are maintained by their owners who curate input from other systems such as publisher, datacentres and institutional systems. ORCID iDs also help to improve the accuracy of information held in, and shared between, systems. As of February 2016, almost 2 million ORCID identifiers have been claimed.

DataCite is a DOI registration agency that provides identifiers for a wide range of academic outputs with a particular focus on data. Datacite works with data centres to assign persistent identifiers to research objects and develop infrastructure that supports simple and effective methods of data citation, discovery, and access. As of February 2016, DataCite have minted over 7 million identifiers.

Authors' Biographies

Tom Demeranville, Senior Technical Officer, THOR Project, ORCID EU.

Tom investigates and develops links with other identifier systems as part of the EC-funded THOR project. Tom has a long history of working with software in the academic sector. Before joining ORCID he was employed by the British Library as a Technical Lead on the ODIN project and prior to that he was a Senior Software Engineer working on federated identity at Eduserv. Tom has a first class honours degree in Software Engineering from the Open University.

Contact: t.demeranville@orcid-eu.org

Josh Brown, Regional Director, Europe, ORCID EU.

Josh directs the operations of ORCID EU. He works with stakeholders across Europe to support understanding and engagement, and promote adoption of ORCID. Before ORCID, Josh was consortium and operations manager for SCOAP3, programme manager for digital infrastructure at Jisc, project officer for SHERPA-LEAP at University College London, and held positions in the library at the University of Brighton and the University of Sussex. He earned an MA in Information Management from the University of Brighton and a BA in Philosophy and English from the University of Sussex.