Geophysical Research Abstracts
Vol. 19, EGU2017-4567, 2017
EGU General Assembly 2017

What is the Value Proposition of Persistent Identifiers?

Jens Klump (1) and Robert Huber (2)

(1) CSIRO, Mineral Resources, Kensington, Australia (jens.klump@csiro.au), (2) MARUM, University of Bremen, Bremen, Germany (rhuber@uni-bremen.de)

Persistent identifiers (PID) are widely used today in scientific communication and documentation. Global unique identification plus persistent resolution of links to referenced digital research objects have been strong selling points for PID Systems as enabling technical infrastructures.

Novel applications of PID Systems in research now go beyond the identification of file based objects such as literature or data sets and include the identification of dynamically changing datasets accessed through web services, physical objects, persons and organisations. But not only do we see more use cases but also a proliferation of identifier systems. An analysis of PID Systems used by 1381 repositories listed in the Registry of Research Data Repositories (re3data.org, status of 14 Dec 2015) showed that many disciplinary data repositories make use of PID that are not among the systems promoted by the libraries and publishers (DOI, PURL, ARK). This indicates that a number of communities have developed their own PID Systems. This begs the question, do we need more identifier systems? What makes their value proposition more appealing than those of already existing systems? On the other hand, some of these new use cases deal with entities outside the digital domain, the original scope of application for PIDs. It is therefore necessary to critically appraise the value propositions of available PID Systems and compare these against the requirements of new use cases for PID.

Undoubtedly, DOI are the most used persistent identifier in scholarly communication. It was originally designed "to link customers with publishers, facilitate electronic commerce, and enable copyright management systems." Today, the DOI system is described as providing "a technical and social infrastructure for the registration and use of persistent interoperable identifiers for use on digital networks". This example shows how value propositions can change over time.

Additional value can be gained by cross-linking between PID Systems, thus allowing new scholarly documentation and evaluation methods such as documenting the track record of researchers in publications and successful funding proposals, apply advanced bibliometric approaches, estimate the output and impact of funding, assess the reuse and subsequent impact of data publications, demonstrate the efficient use of research infrastructures, etc. This recombination of systems raise a series of new expectations and each stakeholder group may have its own vision of the benefits and value proposition of PIDs, which might be in conflict with others.

New PID applications will arise with the application of PID Systems to semantic web technologies and to the Internet of Things, which extend PID applications to beyond digital objects to concepts and things, respectively, raising yet again their own expectations and value propositions. What are we trying to identify? What is the purpose served by identifying it? What are the implications for semantic web technologies? How certain can we be about the identity of an object and its state changes over time (Ship of Theseus Paradox)?

In this presentation we will discuss a number of PID use cases and match these against the value propositions offered by a selection of PID Systems.