THE KNOWLEDGE EXCHANGE PERSISTENT IDENTIFIER PROJECT: BACKGROUND INFORMATION

Maurice Vanderfeesten, Bas Cordewener.

SURF Foundation, Utrecht, The Nederlands e-mail: Vanderfeesten@surf.nl; Cordewener@surf.nl

1. Current PID: status and problem

Worldwide the need has increased to store publications and other information objects with the certainty that they will remain intact and retrievable over (a long) time. This so called 'long term preservation' requires a method to persistently identify information objects.

All over the world initiatives have been deployed to achieve this goal by setting up systems for persistent identifiers.

A variety of solutions have been developed. The existence and use of persistent identifiers is a step forward but presents new problems. Technically, organisational and policy-wise the persistent identifier solutions differ considerably. Many are not really open or interoperable, they have different business models and policy approaches. Due to this, 'link rot' (URL's pointing to no longer existing information) still is a common plague.

For the research and cultural heritage community this is far from ideal: long term preservation and accessibility requirements are not met if there is no interoperability, if the persistence depends on proprietary systems or on third party policies and or commercial drives, which can not be trusted to offer reliable long term references.

2. The PID Challenge

Current persistent identifier systems do not offer the research and cultural heritage community a system that is open and international and is also interoperable with existing persistent identifier solutions. The community is not in command of policy and management of these systems and its specific needs may or may not

be served.

The challenge is to establish a persistent identifier infrastructure which is open, global and interoperable, that is controlled by the research and cultural heritage community and serves their particular needs with regard to long term preservation. These particular needs can be illustrated by the services the community would like to build on top of the persistent identifier systems, opening up possibilities for citation counts, usage statistics and enhanced publications.)

3. The URN-NBN PID Solution

In a number of European countries national persistent identifier solutions are in place, based on the URN-NBN system. These systems have been developed to serve national long term preservation needs. The infrastructure (such as the distribution of unique URN-NBN numbers and long term preservation policies) is managed by national libraries or other national organisations that work for the research and cultural heritage community.

The national systems, which are all based on the URN-NBN approach, work satisfactorily on a national level. Jointly a demonstrator has been built to explore how the national remit of these systems can be elevated to an internationally working system. The URN-NBN system is based on open, existing internet technology standards and the success of this so called Global Resolver indicates that technically there are no real problems to develop and deploy a global persistent identifier system based on URN-NBN.

What needs attention are issues like the organisation and management of an international URN-NBN persistent identifier infrastructure that develops and guards policies and a business model that is sustainable. To ensure true interoperability an effort will be made to update internet standards that define NBNs.

4. Urgency for a global PID solution

4.1. Trust, proprietary ownership, future services

Persistent identification of information objects plays an important role within the research life cycle approach to cultural and scientific digital library applications and the scholarly community. Though on the surface the existence of parallel PID systems may not seem to be a problem, in reality it presents several issues that need to be resolved before long term preservation and persistently identifying re-

search and cultural heritage community information objects will be a trusted, reliable process.

Lack of Trust

Multiple PI mechanisms and many implementers have led to little or no harmonisation on policy about what to identify, how to identify, how to resolve, how long to identify, etc. This blocks researchers and others from trusting and using PIs; there is no coherent information and guidance and without this, the development of a durable network of academic output and cultural heritage and additional services will never be achieved.

Proprietary ownership

Current implementers and/or providers of persistent identifiers have no clear policies about their long term commitments; it is unclear who will guarantee the long term availability of both the persistent identifier and the information object it refers to; it is not transparent how international interoperability will be sought; proprietary drivers will define future decisions and this may not serve the research and cultural heritage community needs.

Future Services

The current situation is too complex and fragmented to build future services on what would really benefit the research and cultural heritage communities, such as worldwide comparison of usage statistics, global citation analysis, portals, enhanced publications, metadata interoperability facilities, global semantic thesauri, detection of duplicates, etc.

In this context it is currently difficult for researchers and applications to work with NBNs because they depend on the availability of a nationally available resolver, which does not yet have the level of trust required, and only offers a particular, nationally defined persistence. These national NBN services are relatively unknown.

4.2 The URN-NBN option

It is urgent to build a trusted and durable infrastructure for scholarly communication and cultural heritage that provides permanent and transparent access to both research and cultural heritage resources on the Internet.

Such a durable infrastructure requires an open technology approach, a trusted

framework of management and policies directed at the research and cultural heritage community, and clear information to all involved about opportunities and procedures.

This durable infrastructure can be realised building on the URN-NBN concept, in particular on the basis of

- previous satisfactory experience with URN-NBN in national contexts,
- the successful building of a global URN-NBN resolver demonstrator,
- the willingness of national representatives of research and cultural heritage communities to engage in a shared URN-NBN infrastructure,
- URN-NBN being fully based on open, interoperability standards and being able to resolve and harmonise other existing persistent identifiers.

This situation opens the opportunity for a project that will

- bring immediate benefits to URN-NBN using partners,
- establish a mature international URN-NBN infrastructure,
- explore URN-NBN potential to harmonise current PID diversity.

June 2009
Printed on demand
by "Nuova Cultura"
www.nuovacultura.it

Book orders: ordini@nuovacultura.it