

Annex Use of Persistent Identifiers

Persistent Identifiers for web resources are needed to create a stable and reliable infrastructure. This does not concern technicalities, but mainly agreements on organisational level.

DRIVER Guidelines could make some recommendations on the implementation for repository managers. This is based on the [Report on Persistent Identifiers of the PILIN project](#).

An implementation plan has been provided below.

It should be made clear how this fits in with oai_dc exchange of metadata

In the era of paper the International Standard Book Number (ISBN), a unique, numerical commercial book identifier, was developed. Each edition and variation (except reprinting) of a book is given an ISBN. In the digital age, there is a growing need for such a unique, numerical, identifier for digital publications as well. Moreover, not just for publications, but for all kinds of digital objects.

On the Internet, we consider the URL as the identifier of a digital object. However, we are all familiar with broken or dead links that point to web pages that are permanently unavailable.

An URL might change overtime, due to server migrations and other technical reasons. With undesired consequences for links and citations within scholarly communication.

Therefore a 'persistent identifier' is needed with which a digital object is permanently associated. This persistent identification number always refers to the digital object to which it has been assigned, regardless of the underlying locator technology (at the moment these are web addresses; in the future, however, an object's location may be completely different).

In several countries, a system for such a persistent identifier has been developed and 'national resolvers' have been set up. A resolver is a transformation and redirection service, transforms a string of characters to an URL, and is hosted by a national organisation. Common identifiers in the case of scholarly communication are DOI, Handle and URN:NBN. In case of DOI and Handle the resolution mechanism is located in the US at [CNRI](#). In case of URN:NBN resolution mechanisms are hosted by a national organisation, often this is done by the National Library.

Every digital object is assigned a number that represents that object forever. Even if technology moves on, the national organisation will ensure that the documents can be read. But the documents must be traceable as well. The Persistent Identifier ensures that it can be located. A stable information infrastructure makes research citations a lot more reliable.

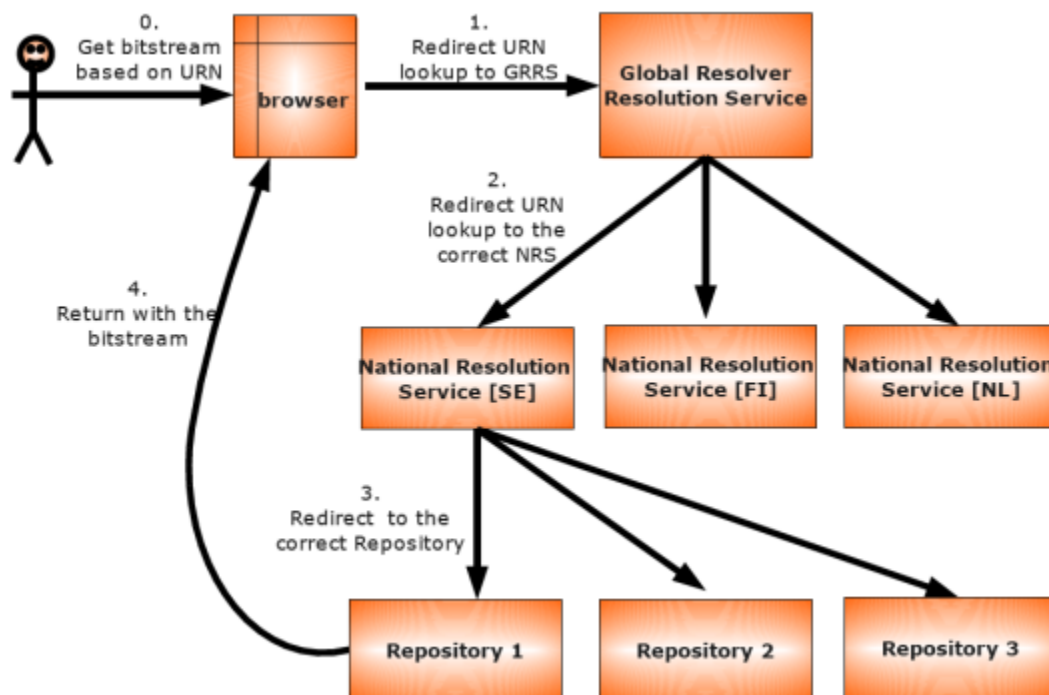
Currently the URN:NBN and the Handle are popular ways for Persistent Identifiers. Since the URN:NBN namespaces are distributed in a controlled manner, we would expect it will be recognised as authoritative as the DOI has as a reputation.

The differences between Persistent Identifiers are described by Hans-Werner Hilse and Jochen Kothe in [Implementing Persistent Identifiers \[Hilse\]](#). There is also an article [Persistent Identifiers: Considering the Options](#) in Ariadne, issue 56 by Emma Tonkin [\[Tonkin\]](#).

Using Persistent Identifiers involves an obligation for the repositories to sustain persistence of the Identifier over a long period of time! This persistence can be guaranteed in so called "trusted repositories" with the appropriate certification. See chapter [Annex Use of Quality Labels](#) page. for more information see <http://www.persistent-identifier.de> and <https://www.pilin.net.au/>

The Scandinavian countries, Germany, the Czech Republic and the Netherlands are using URN:NBN. The main reason for choosing urns is because it is an internet standard that is future proof. The only drawback now is that a urn is not actionable without using an http resolution address as a prefix. Further work is still needed to be done to [integrate URN in the DNS system](#) by using [NAPTR records](#) that is also used for VOIP phone calls.

Recently Norway, Sweden, Finland, and the Netherlands have come to a promising proposal for a Global Resolver of Persistent Identifiers (URN: NBN). In cooperation with representatives of the Hopkins and Berkeley Universities (US) a working [proof of concept](#) Global Resolution Proof of Concept: <http://www.surfgroepen/sites/surfshare/public/software/pihandler> of a global resolver (GRRS) has been developed. This GRRS integrates four different national resolvers into one global resolver. The GSRS ([n2t.info](#)) receives the Identifier from a browser plug-in and redirects the browser to the appropriate national resolver where the browser again is redirected to the current location of the web resource. The architecture of this multi-system process is depicted below.



Implementation plan on using URN_NBN Persistent Identifiers

First of all we would like to say that the persistency of Identifiers and web resources is not about the technology one uses, but about organisation and sustainable business models. For more information about Persistent Identifier policies take a look at the successful [Persistent Identifier Linking INfrastructure \(PILIN\)](#) project in Australia that is part of the [ARROW](#) project.

To setup a persistent Identifier program based on [National Bibliographic Numbers \(NBN\)](#) URN identifiers and a resolver one needs to take the following steps:

1. **Work group:** Create a work group that manages all the technical and organisational details of such project. Also think about the syntax that is going to be used. For example `urn:nbn:{country}:{sub-namespace}:{repositoryid}-{localid}`. Country is the short name of the country, sub-namespace represents web resources that come from the repositories, repositoryid is a two digit representation of the repository and local id is the Identifier generated at the repository. This can for example result in the following Identifier for one publication `urn:nbn:ie:ui:21-1234/5678`.
2. **Formalities:** Since the `urn:nbn:ie` namespace is by default claimed by the National Library, one has to arrange an agreement with the National Library to use a sub-namespace for scientific material. This name should be short and have no semantic meaning. For example `urn:nbn:ie:ui`, or `urn:nbn:ie:oa`, or `urn:nbn:ie:sp`.
3. **Registration Agency:** Create a registry in which repositories are given a short random number of two digits. This will create a sub-namespace in which a repository autonomously can distribute Persistent Identifiers for their publications. For example Trinity College Dublin (TCD) is registered as 21. The namespace for TCD to operate in will be `urn:nbn:ie:ui:21`.
4. **Implementation at local level:** Each repository must generate Persistent Identifiers for each publication within their namespace that is provided and store this identifier in the database record. For example TCD can use existing identifiers to add after their namespace followed by a dash. In case TCD uses handle, the Identifier for one publication could look like the following `urn:nbn:ie:ui:21-1234/5678`. In case TCD uses database numbers `urn:nbn:ie:ui:21-15874`. (Make sure to store the identifier and not generate them on-the-fly. In case of database migrations these numbers might change and persistency is lost.)
5. **Transport of identifiers and URL's:** Each repository must generate a DIDL package in which the URN and URL are included. See the MPEG-21 DIDL section in the main report.
6. **National Resolution Service:** A national resolver can be made by harvesting the DIDL packages from each repository where the URL and URL bindings are extracted and stored. A web location must be created where the user or machine can go to for resolution of the identifier. For example <http://resolver.ie> where the user can insert an identifier and receive the current location of the web resource. For example <http://resolver.ie/urn:nbn:ie:ui:21-1234/5678> resolved to <http://repository.tcd.ie/1234/5678>

References

[Hilse]	Hilse, H., Kothe, J., Implementing Persistent Identifiers , KNAW, http://www.knaw.nl/ecpa/publ/pdf/2732.pdf
[Tonkin]	Tonkin, E., Persistent Identifiers: Considering the Options, Ariadne, issue 56, http://www.ariadne.ac.uk/issue56/tonkin/
CNRI	http://www.cnri.reston.va.us/