Persistent Identifiers distributed system for Cultural Heritage digital objects

Emanuele Bellini *, Chiara Cirinnà *, Maurizio Lunghi *,

Ernesto Damiani, Cristiano Fugazza****

*Fondazione Rinascimento Digitale
Via Bufalini, 6, 50100, Florence , ITALY
{bellini,cirinna,lunghi}@rinascimento-digitale.it

**Information Technology Dept.
University of Milan
Via Bramante, 65, I-26013 Crema (CR), ITALY
{damiani,fugazza}@dti.unimi.it

Abstract

In this paper, we present a prototype with a novel resolution architecture for an URN based Persistent Identifiers (PI) system in Italy. We describe a distribute approach for implementing the NBN namespace system and illustrate the solutions adopted for the assignment and resolution of the identifiers with the hierarchical and peer-to-peer request forwarding.
Starting from the core motivations for 'persistent identifiers' for digital objects, we draw up a state of art of PI technologies, standards and initiatives, like other NBN implementations. The prototype is still under development and we present the next steps, in particular we describe the interoperability perspective that already partially foresees the NBN prototype.

Introduction

Persistent identification of Internet resources is an important issue within the life cycle approach to cultural and scientific digital library applications, not only to identify a resource in a trustable and granted way, but also to guarantee continuous access to it. It is well-known that Internet resources have a short average life; their identification and persistent location poses complex challenges affecting both technological and organizational issues, involving access and citation of cultural and scientific resources. The use of URLs can not be considered a reliable approach due to the structural instability of links (ex. domains no longer available) and related resources relocation or updating. The current use of the URL approach increases the risk of losing cultural documents or under-using available cultural collections.

The issue is more organisational than technical. There are already some initiatives aiming at stabilising Internet addresses, for example setting up a central registry with a stable reference/name of a resource with a redirect to the actual URL. But for us this is simply not enough. In the Cultural Heritage (CH) domain it is essential not only to identify a resource but also to guarantee authenticity, credibility and continuous access to it.

In synthesis, a first essential component in realising a 'long-term availability' is the use of Persistent Identifiers (PI) in order to solve the problem of univocal identification and **reliable locator** of Internet resources. But another key component to implement a PI service is the **credibility** and long term sustainability of the **Registration Authority,** the institution that stands security for the maintenance of the PI-URL association register, and granting for the resource authenticity, completeness and the content accessibility. Another element to be taken in consideration creating a name space for any type of resources is the level of **service** and '**granularity'** that the identifiers are requested by the specific user application.

Persistent identifiers solution

A trustworthy solution in the CH is to

associate a Persistent Identifier (PI) to a digital resource certifying in some way its content authenticity, provenance, managing rights, and providing an actual locator.
Persistence refers to the permanent lifetime of an identifier. It is not possible to reassign the PI to other resources or to delete it. That is, the PI will be globally unique forever, and may well be used as a resource's reference far beyond the lifetime of the identified resource or the naming authority involved. Persistence is evidently a specific matter in a cultural institution's service or policy. The only guarantee of the usefulness and persistence of identifier systems is the commitment shown by the organizations who assign, manage, and resolve the identifiers.
Each PI system foresees the existence of a Registration Authority (RA). The RA is an independent authority that assigns names and guarantees their uniqueness and persistence. Finally, the service tailored on user needs, a naming resolution service corresponds to every naming authority and carries out the name resolution.

These are the main steps to be performed in order to implement a PI system:

1) Selection of resources that need a PI and define the level of granularity requested by the user application.
2) Identification of a RA suitable/trustable for the digital content and the specific user application. A business model sustainable must be defined.
3) Definition of the level of service for resolution of names, in particular the resource info data presented, the rights and access modalities.
4) Execution of resource name creation and assignment of one or more URLs in the system register.
5) Execution of a resolution service for couples PI-URL.
6) Maintenance of the register that associates PI-URL and guarantees of continuous access to the resources.

The first three steps are prerogative of each cultural institution or user application manager, whereas the steps thereafter can be delegated to other authorities, in order to guarantee better economic and functional sustainability of the service.

State of the Art

At present some technological solutions (e.g. DOI, ARK, Handle system, URN) have been already developed but no general agreement has been reached among the different user communities so far: this scenario shows that it is not viable to impose a unique PI technology. Moreover the granularity, that refers to the level of detail at which persistent identifiers need to be assigned, is widely different in each user application sector.
Among existing standards for PIs, the more relevant seam to be the following: Uniform Resource Names (URN), Life Science Identifiers (LSID), Persistent URL (PURL), Archival Resource Key (ARK), Handle System with its Digital Object Identifier (DOI) implementation, and the Library of Congress Control Number (LCCN).
URN is a key standard issued by the IETF and experts are promoting that as a meta name-space in order to include other identification systems. PURL is simply a redirect-table of URLs and it's up to the system-manager implement some policies for authenticity, rights, trustability. LCCN is something similar but with a credible policy for trustability and stability of identifiers. The DOI system, is a business-oriented solution widely adopted by the publishing industry and that provides administrative tools and a DRM System. ARK provides peculiar functionalities that are not featured by the other PI schemata, e.g., the capability of separating the univocal identifier assigned to a resource from the potentially multiple addresses that may act as a proxy to the final resource. Furthermore, we may also find multiple, proprietary implementations for a given schema: the URN-based schema grounded on NBNs has been registered and adopted by the Nordic Metadata Projects but is being separately implemented by individual systems with no reference implementation enabling coordination of information sources.

The Uniform Resource Name Approach

The purpose of a Uniform Resource Name (URN – RFC1737) is to provide a globally unique, stable, location-independent resource identifier which can be used for identification, for access to resource characteristics or for access to the resource itself. The URN specification is part of the IETF family of specifications encompassed by the Uniform Resource Identifier (URI) framework. This framework also includes URLs, which specify both a protocol and a location in order to give access to resources on the web. IANA (Internet Assigned Numbers Authority) is the

Registration Authority (RA) for URN namespaces. URNs are designed to enable heterogeneous namespaces mapping and currently, experts are promoting this standard as a common level of integration/interoperability with other 'traditional' identification systems like ISBN-ISSN-SICI (see RFC2288, RFC3044, RFC3187).

Unlike URLs, URNs are not directly actionable (browsers generally do not know what to do with a URN) because they have no associated global infrastructure that enables resolution (such as the DNS supporting URL). Although several implementations have been made, each proposing its own means for resolution through the use of plug-ins or proxy servers, an infrastructure that enables large scale resolution has not been implemented. Moreover, each URN name-domain is isolated from other systems and, in particular, the resolution service is specific (and different) for each domain.

NBN namespace and on-going projects

The National Bibliographic Number (NBN – RFC3188) is a namespace used by National Libraries and based on the standard URN by the IETF. The NBN namespace, as a Namespace Identifier (NID), has been registered and adopted by the Nordic Metadata Projects on request of the CDNL and CENL.

The RFC 3188 says:

> 'The NBN is a generic name referring to a group of identifier systems utilized by the national libraries and only by them for identification of deposited publication which lack an identifier, or to descriptive metadata (cataloguing) that describes the resources'.

Each National Library uses its own NBN string independently and separately implemented by individual systems with no coordination and no common formats with other national libraries. In fact, several national libraries have developed their own NBN systems for national and international research projects; several implementations are currently in use, each with different metadata descriptions or granularity levels. An example is the DIVA project at the Uppsala University Library in Sweden, where documents published in the DIVA-Portal have a unique identifier. In cooperation with the Royal Library of Sweden, they implemented an URN-NBN system. One can access every document registered on the DIVA system from the NBN resolver at the Royal Library, whether it may be located at its originating institution or at the Royal Library archive. A similar example is the EPICUR Project at the Deutsche Nationalbibliothek. The aim of the project is to enhance the existing URN-NBN system in Germany for online theses and other types of resources.

There are some important initiatives at European level like the TEL project that it is in the process of implementing a unique system based on NBN namespace within the European Digital Library (EDL). The adoption of NBN identifiers is needed for implementing the 'National Libraries Resolver Discovery Service' as described in the

CENL Task Force on Persistent Identifiers, Report 2007.

The NBN Project in Italy

The project, funded by the Fondazione Rinascimento Digitale (FRD) and developed together with the National Library in Florence (BNCF), the University of Milan (UNIMI), and the University consortium (CILEA), has developed a prototype for a national register of digital cultural resources. The first phase of the project has already been completed and the first results are available; future objectives are defined looking for international cooperation. The NBN project is based on a 'trusted digital repository' installed within another project jointly developed by the FRD and BNCF.

First phase objectives

The main objectives of the project first phase have been the following:

- to create a national stable and certified register of digital objects in use by cultural and educational institutions;
- to allow an easier and wider access to the digital resources produced by Italian cultural institutions, including material digitised or not yet published;
- to encourage the adoption of long term preservation policies and make costs and responsibilities for the service sustainable;
- to test a new technology based on URN but upgraded in its architecture with distribution of responsibility for names management;

- to create some redundant mechanisms both for duplication of name-registers and in some cases also for the digital resources themselves.

Second phase objectives

The main objectives of the project second phase will be the following:

- to extend as much as possible the adoption of the NBN technology and the user network in Italy;
- to reinforce the peer-to-peer resolution service and the robustness of the network for direct access to digital resources;
- to develop a protocol for inter-domains (e.g., NBN Italy and NBN Germany, or NBN Italy and DOI) resolution service with a common format of info-data and a friendly user interface.

Hence in the CH context, it is necessary to implement a service for URN assignment and resolution on the national level (managed by the National Library of Florence - BNCF), based on NBN. The decision to utilise the NBN is due to the fact that it is a namespace for the exclusive use of national libraries (every country has registered a sub-namespace at the Library of Congress: for Italy: NBN:IT ISO 3166); this guarantees the presence of the requisites of stability and permanence necessary for an institution that intends to manage a PI service.
The project has developed a prototype that, independently from the content management systems of the single cultural institutions, realises a national register of persistent identifiers for the digital cultural objects on the Internet, and experiments a service of resolution and access to these resources by inserting several elements of novelty in the system's architecture and functionality with respect to the technological solutions currently proposed or under development.
This solution is conceived especially for those resources that do not have any type of identification (i.e. doctoral theses, digitisation of antique books, etc), but in perspective, it can also be extended to unifying all digital cultural resources under a single code, even those are already identified by codes like ISBN, ISSN, SICI, or DOI. The identified resources will thus be able to reside on the system of the cultural institutions that have the rights to manage their sub-domains, and in the legal deposit system.
Therefore the expected impact of the project will be to extend, as thoroughly as possible on the national level, the adoption of this technology of stable addressing of the Internet resources in the sphere of culture and education, valorising the scientific and cultural production of the Italian institutions and improving their impact on the public, including that of works that are either little-known or out of print.

A Distributed Approach

A PI distributed system foresees that the responsibility of generation and resolution can be delegated to other institutions called sub-naming authorities, who manage a portion of the name domain/space.
The Italian prototype implements a new PI architecture: the approach is based on URN/NBN, with additional features and solutions recalling the DNS architecture. The prototype defines a hierarchical distributed system, in order to face the criticality of a centralised system and to reduce the high costs of management for a unique resolution service preserving the authoritative control. The project foresees a distributed authority and responsibility for the creation and resolution of names and redefines the central point role, from a unique name generation/resolution access point to an identifier validator and resolution request router. The central node (BNCF) manages the entire domain NBN:IT, but delegates some second-level agencies to manage sub-domains (e.g., NBN:IT:FRD) both in terms of generation and resolution of names.
In this approach the level of resolution trustability increases if the number of institutions joining the network grows up, because, differently from other approaches, in this pilot there is not a single URL for accessing the NBN resolution service. Every node resolves every NBN item generated inside the sub-namespace IT.
The responsibility distribution joined, in some cases, with duplication of data (names+resources) help increase robustness and performance of the system.

NBN System Architecture – Elements &Functions

The architecture of this NBN system is carried out on two levels, with five elements and four basic functions.

ELEMENTS

1) central node (BNCF)

2) sub-domains register
3) second level agencies (cultural institutions)
4) NBN sub-domain register
5) NBN central register

MAIN FUNCTIONS

1) creating a sub-domain
2) generating a name
3) updating the NBN registers
4) resolving a query for a name

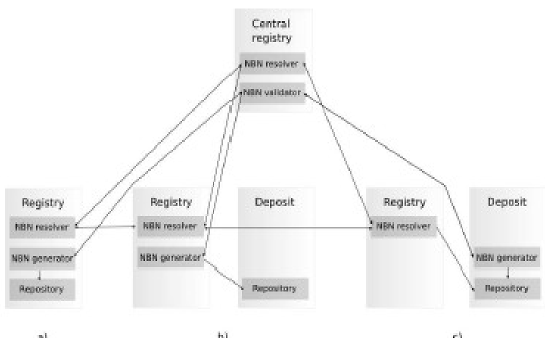


Figure 1 – NBN architecture

Central point
The architecture identifies a central point, located at the National Library of Florence (BNCF) the Registration Authority for the Italian NBN domain, and some second-level institutions. The central node can generate names and sub-domains; it can resolve a user-query directly or redirect it to the appropriate second level agency.
The central node acts in some cases as ‘legal deposit’ archiving also the digital resources.
The system is designed to separate the resolution service from the deposit of the resources.

Sub-domains register
Each second level node is identified by a sub-namespace expressed through the NBN name (for example NBN:IT:BNCR:xxx-xxxx for the National Library of Rome). This registry holds the associations between the sub-namespaces and the base URL of the second level registered institutions. This register is located in the central node for the harvesting function, as well as in the second level nodes for allowing the peer-to-peer resolution process.

Second level agencies
The second level nodes manage their sub-domains, like a DNS, generating names for resources on user demand, keeping a sub-domain register updated with all the associations NBN-URL for their sub-domain names. Most of them offer also a resolver service with a web interface: for names belonging to their sub-domain they are able to solve and provide direct access to the digital resources, for other names they ask the central node to resolve the query or try in peer-to-peer with other second level agencies.

NBN sub-domain register
It is specific for each sub-domain and list the names registered by the second level nodes with all the associations NBN-URL for their sub-domain names.

NBN central register
The central node harvests in OAI-PMH each sub-domain register to check the new names, avoid duplication of names for the same resources, and updated the central register with all the associations NBN-URL for the entire Italian domain. In some cases, it may also have a copy of the digital resource itself, creating a double URL association for that name.
In order to avoid the management costs, the register does not include the descriptive metadata of resources, but only the administrative metadata for managing NBN name’s lifecycle and an external pointer to authoritative metadata belonging to existing institutional repositories.

Creating a sub-domain
The BNCF can generate a sub-domain for any authorised institution providing it with a prefix like NBN:IT:FRD and include this in the sub-domain register that will be also redistributed to all the other second level nodes. The central node checks periodically the status of the second level agencies and the new names generated.

Generating a name
A name can be generated on user demand, by the central node or by any of the second level agencies, but when a new NBN is generated by institutions, it is not immediately resolvable: an answer is expected from the central point to check uniqueness of name-resource combination. Names are not reusable or changeable. The NBN central register lists all the names within the NBN:IT domain.

Updating the NBN registers

The central point is composed of the central register where there are stored all NBN names generated from any second level institution, a check module of the NBN harvested and a sub-domains register with the URLs of all the second level agencies. We have already seen that the central node is responsible for updating and distributing the sub-domains register. About the names, the central point harvests periodically the NBN records from the second level nodes, then an automated process verifies if the NBNs harvested are correct (see Registries syncronization). Finally the central point sends an answer where are pointed out the NBNs that are not correct or a simple confirmation message if there are no problems, and of course it updates the NBN central register.

Resolving a query for a name

Any second level agency may have installed the resolver service through a simple web page for any name of the NBN:IT domain. If the name requested by the user belongs to the same sub-domain the second level node resolves directly the query, otherwise asks the central node or other agencies in peer-to-peer. The answer is both some info-data and the direct link if access-rights are available.
In future, the same web page will be able to resolve also names belonging to other NBNs or to DOI.

Other Functionalities in deep

NBN administrative metadata format

```
<xs:schema>
<xs:element name="nbn-record">
<xs:complexType>
<xs:sequence>
<xs:element name="URI" minOccurs="1"></xs:element>
<xs:element name="URL" minOccurs="1" maxOccurs="unbounded"></xs:element>
<xs:element name="metadataURL" minOccurs="0">
</xs:element>
<xs:element name="MD5" minOccurs="1"></xs:element>
<xs:element name="creationDate" type="xs:dateTime" minOccurs="1"/>
<xs:element name="lastModified" type="xs:dateTime" minOccurs="1"/>
<xs:element name="status"></xs:element>
<xs:element name="event" minOccurs="0"></xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
</xs:schema>
```

The field creationDate and modifiedDate are used by harvesting engine in order to perform a differential haversting. In particular the MD5 is an hash field calculated for the physical digital object with the MD5 algorithm . This field is very important for the central point because allows to check if a resource has multiple identifiers. The field status and event are used to track the NBN life cycle as described here below.

NBN life cycle

The project foresees to track each event that may affect the NBN identifier record. There are several “actions” that are managed, like NBN creation or NBN record update. The tracking of update action is important when the resources change their location on the net and consequently change their URL. Another important update action takes place when a new URL is added to for the multi URN-URLs association. The details on the life cycle of NBN identifiers are rendered in Fig 1 as a finite state automa:

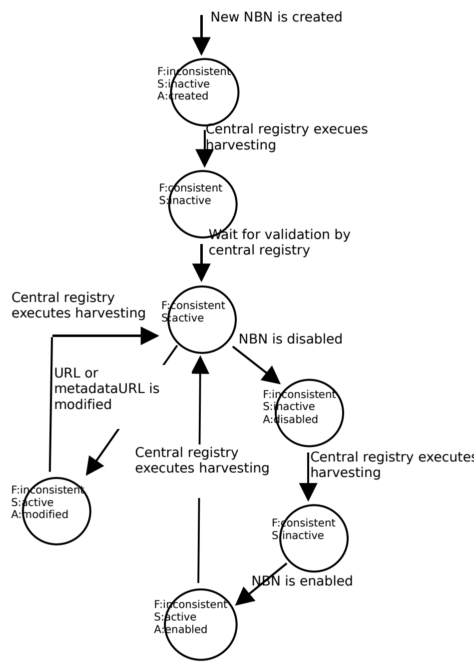


Figure 2 – NBN life-cycle state automa

The distinct states can be singled out by using three variables:

– Flag F, whose value is either ’inconsistent’ or ’consistent’, that determines whether or not the central register should harvest the record associated with the NBN because it has changed.
– Status S, whose value is either ’active’ or ’inactive’, that indicates whether registers should resolve the NBN, that is, if the resource associated with the NBN is currently available.
– Action A is an additional variable indicating to the central register, during metadata harvesting, the particular operation that has been carried out (allowed values are ’created’, ’enabled’, ’disabled’, and ’modified’).

Registries Synchronization

The architecture foresees the synchronization between central register and second level registries through OAI-PMH protocol. The central node manages the register of sub-namespaces necessary for harvesting of metadata from second level nodes.
This process has 3 steps:

1) Harvesting NBN records
2) Check NBN records
3) Answer to second level register

Harvesting NBN records
The first step is a differential harvesting of NBN records from second level registries. Only the new NBN or NBN records affected by an update will be harvested.

Check NBN records
The second step is the check the data consistency.
Case of alert:

a) Different NBN and same MD5
An identifier must be assigned to a single resource. If there are other copies of the same resource, the system manages the multi association URN-URLs. The institution that has tried to generate an NBN for a resource that has already an NBN receives a message indicating the right name to be used for that resource.
b) Identical NBNs and different MD5
c) Identical NBNs, MD5 and lastModified
These two cases are errors that could happen for many reasons. The prototype sends an alert to the responsible of the last harvested NBN, in order to check and face the problem.

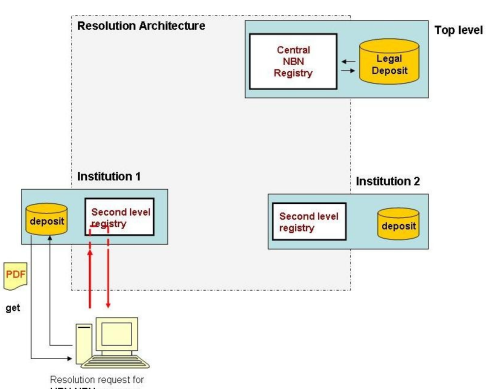
Answer to second level node
The third step is to send the check results to the institution in order to manage the inconsistency. If the problem is the a) case, the institution should disable its NBN identifiers (that will be enabled for another registration). Another email is sent to the owner of the NBN-URL-MD5 first- registered with the new URL of the copy of the same resources included. The owner adds this new URL in the field URL of their NBN registry. The next harvesting and identifier check will enable this multiple resolutions.
If the problems are cases b) or c), other activities should be planned like software debug or check for a human error.

NBN Resolution Process

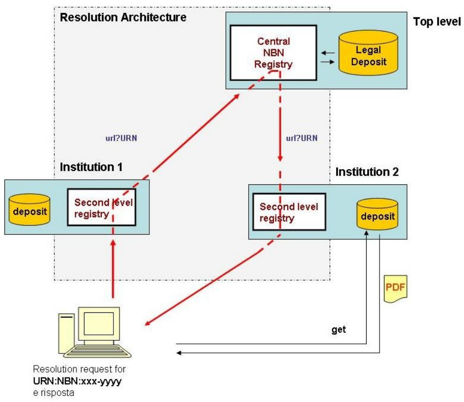
The name resolution request can be submitted by the user to any resolution service of the second level nodes. If the sub-namespace identifies the institution to which the request is submitted, the answers is given directly, otherwise the central registry will be invoked to redirect the resolution request to its appropriate second level node.
This architecture increases the robustness of the service and also foresees a peer-to-peer resolution between the second level institutions, in order to maintain the resolution infrastructure operational, even if the central service is not available.

The cases of interaction are the following:

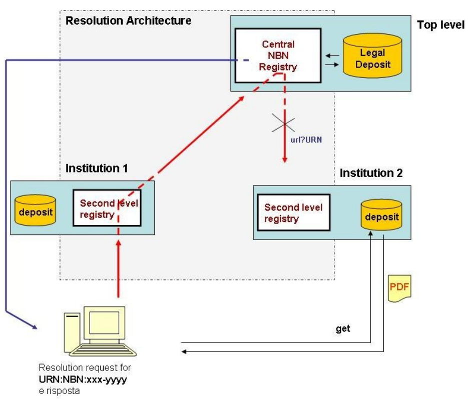
a) the sub-namespace identifies the institution to which the request is submitted. If the resolution request of this name URN:NBN:IT:FRD:xxx-xxxxx is submitted to FRD resolution service, the answers is given directly.



b) the sub-namespace does not identify the institution to which the request is submitted. If the resolution request of this URN:NBN:IT:FRD:xxx-xxxxx is submitted to BNCF resolution service, the central registry is invoked to redirect the resolution request to its appropriate second level node.



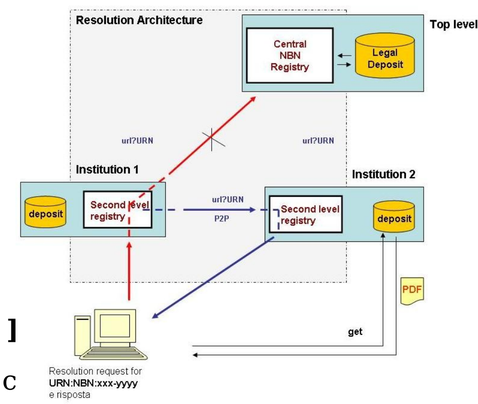
c) the ... the instituti... itted and the ... does not wo... s on behalf of the second level node.



d) the sub-namespace does not identify the institution to
which the request is submitted and the central registry does not work: the second level node activates the peer resolution. This is a one of the key features of the entire architecture. This solution can be used as a resolution safe mode or could be selected by a load balancing strategy.
In fact a specific load balancing service could decide to foreword the resolution request to the peer to peer channel or to the hierarchical resolution process every time.

Peer to peer resolution process

The developed architecture foresees a peer resolution of NBN identifiers. This solution is useful when the central resolution service for any reasons does not work. This feature is necessary because the hierarchical approach has still a single point of failure for a full resolution service. When the central point dos not work, every second level institutions are able to resolve their NBN identifiers only without a peer resolution system. The trustability of peer resolution depends on what synchronization strategies are adopted to line up the second level sub-namespace registry with central point sub-namespace registry. Each second level point has a copy of the sub-namespace registry of the central point. The second level resolution service is able to recognize the sub namespace of the NBN string and foreword this resolution request to the appropriate second level institution, using the sub-namespace registry. The use of peer-to-peer resolution as a back up service of the hierarchical resolution is a choice. In fact is possible to set the peer resolution as primary and call the central register only if there is no answer from peer or implementing a load balancing service as described above.



]

The c ... bility functionalities with other namesystems, included other NBN systems, as well as the DOI system. The PI systems are thought as autonomous systems. The NBN project has designed the central node as a gateway to forward towards other domain (NBN:DE, DOI, ARK) the resolution request of other NBN namespace identifiers. This approach is a first step for a wider interoperability project among different PI domains for a common resolution service.
This function is under development.

Ongoing research activity

The FRD in conjunction with mEDRA (European DOI Registration Authority), CINECA, CILEA, the University of Milan, the central Library of National Research Council (CNR) in Italy are developing a common base resolutions service with DOI, in order to realise a full interoperability with these two identification systems. The approach follows also the CENL recommendations as the 'last resort resolutions' of DOI by NBN. Another important development of the pilot is to establish a common resolver service with other NBN systems in other countries.

Outcome and expected impact

The expected impact is not only a great improvement in the quality of the web coverage of European cultural resources and the reduction of costs and efforts needed to maintain a stable reference of Internet resources, but also a general increasing of credibility and trustability for digital libraries, by promoting the use of digital contents in different user sectors and applications. In particular:

- Persistent identifiers and NBN promotion
- Prototype development and technology evaluation
- Open Source technologies promotion
- Digital preservation development
- Preservation of the minor literature
- Access to resources of difficult or impossible search

References

J. Hakala. Using national bibliography numbers as uniform resource names, 2001. (RFC3188) http://www.ietf.org/rfc/rfc3188.txt.

J. Kunze. The ARK Persistent Identifier Scheme. Internet Draft, 2007. http://tools.ietf.org/html/draft-kunze-ark-14.

C. Lagoze and H. V. de Sompel. The Open Archives Initiative Protocol for Metadata Harvesting, version 2.0. Technical report, Open Archives Initative, 2002. http://www.openarchives.org/OAI/openarchivesprotocol.html.

Norman Paskin. *Digital Object Identifiers*. Inf. Serv. Use, 22(2-3):97–112, 2002.

Kathrin Schroeder. *Persistent Identification for the Permanent Referencing of Digital Resources* - The Activities of the EPICUR Project Enhanced Uniform Resource Name URN Management at Die Deutsche Bibliothek. The Serials Librarian, 49:75–87(13), 5 January 2006.

Sam X. Sun. Internationalization of the Handle System - *A persistent Global Name Service*. 1998. http://citeseer.ist.psu.edu/sun98internationalization.html.

DCC Workshop on Persistent Identifiers
30 June – 1 July 2005 Wolfson Medical Building, University of Glasgow
http://www.dcc.ac.uk/events/pi-2005/

ERPANET workshop Persistent Identifers
Thursday 17th - Friday 18th June 2004-University College Cork, Cork, Ireland
http://www.erpanet.org/events/2004/cork/index.php

Object Management Group. Life Sciences Identifiers final adopted specification, 2004. http://www.omg.org/cgi-bin/doc?dtc/04-05-01.

Dublin Core Metadata Initiative. Dublin Core Metadata Element Set, Version 1.1. http://dublincore.org/documents/dces/.

H.-W. Hilse, J. Kothe *Implementing Persistent Identifiers: overview of concepts, guidelines and recommendations*,
2006, ix+57 pp. 90-6984-508-3
http://www.knaw.nl/ecpa/publ/pdf/2732.pdf

E. Bellini, C. Cirinnà, M.Lunghi, 2008, *Persistent Identifiers for Cultrual Heritage*, DigitalPreservationEurope Briefing Paper
http://www.digitalpreservationeurope.eu/publications/briefs/persistent_identifiers.pdf

E. Bellini, M.Lunghi, E. Damiani, C. Fugazza, 2008, *Semantics-aware Resolution of Multi-part Persistent Identifiers*,
WCKS 2008 conference.

CENL Task Force on Persistent Identifiers, Report 2007
http://www.nlib.ee/cenl/docs/CENL_Taskforce_PI_Report_2006.pdf.

National Library of Australia, PADI (Perserving Access to Digital Infornmation) Persistent Identifiers , 2002
http://www.nla.gov.au/padi/topics/36.html#article

Relationship Between URNs, Handles, and PURLs Library of Congress, National Digital Library Program
http://lcweb2.loc.gov/ammem/award/docs/PURL-handle.html

Corporation for National Research Initiatives *Handle System (Overview)*
http://www.handle.net/overviews/overview.html

Coyle, Karen Identifiers : *Unique, Persistent, Global*
http://dx.doi.org/10.1016/j.acalib.2006.04.004
Journal of Academic Librarianship, Volume 32, Issue 4, (July 2006), p. 347-450

Bermes, Emmanuelle, International Preservation News, Vol 40 December 2006, pp 23-26 *Persistent Identifiers for Digital Resources: The experience of the National Library of France*
http://www.ifla.org/VI/4/news/ipnn40.pdf

Andersson, Stefan; Hansson, Peter; Klosa, Uwe; Muller, Eva; Siira, Erik Using XML for Long-term Preservation : Experiences from the DiVA Project

David Giaretta, Issue 1, Volume 2 | 2007
The CASPAR Approach to Digital Preservation
The International Journal of Digital Curation

Sollins, Karen *Architectural Principles of Uniform Resource Name Resolution* (RFC 2276)
http://www.ietf.org/rfc/rfc2276.txt

Masinter, Larry; Sollins, Karen *Functional Requirements for Uniform Resource Names* (RFC 1737) http://www.ietf.org/rfc/rfc1737.txt

Moats, Ryan 1997*URN Syntax* (RFC 2141)
http://www.ietf.org/rfc/rfc2141.txt