

The application of Persistent Identifiers as one approach to ensure guarantee long-term availability of Online-Theses

The established Uniform Resource Name (URN) -Management at Die Deutsche Bibliothek

Kathrin Schroeder

Die Deutsche Bibliothek

schroeder@dbf.ddb.de

Adickesallee 1, D-60322 Frankfurt am Main, Germany

www.ddb.de

Keywords: Persistent Identifier, Uniform Resource Name, long-term availability, permanent referencing, Die Deutsche Bibliothek

Abstract

Long-term availability of digital objects e.g. online dissertations embedded in a distributed heterogeneous information environment causes new challenges. This applies particularly to national libraries with a legal mandate for collecting, archiving and indexing national cultural heritage. Several aspects must be considered in order to guarantee long-term access to the digital resources e.g. the preservation of the bitstream of digital objects, realizing migration, and emulation strategies. An essential component to achieve long-term availability is the use of Persistent Identifiers (PI) that solve the problems of unequivocal resource identification and reliable addressing.

This article discusses the implementation of URNs at the German National Library (Die Deutsche Bibliothek) within the international name space of national libraries "National Bibliography Number" (NBN). Following, the practical experience of URN management at Die Deutsche Bibliothek will be presented.

Preface

One of the recommendations of the report of "Trusted Digital Repositories: Attributes and Responsibilities - an RLG-OCLC Report"[1] is the use of Persistent Identifiers (PI) within digital depot SYSTEMs in order to provide a long-term access to digital objects in conjunction with an intuitive usable mechanism that enables the enduser to get a desired resource. To achieve these objectives two different aspects must be taken into account within a PI SYSTEM. The first aspect concerns the unequivocal resource identification, the second aspect is reliable addressing. Both aspects are discussed below.

The existing structural variety of digital objects creates new challenges for an efficient design of identification SYSTEMs. Examples are multimedia applications, hierarchical, and modular digital objects. However powerful identification SYSTEMs are not sufficient in order to guarantee long-term availability of digital resources without a method for reliable access to digital resources.

At present there are several mechanisms to maintain URLs relatively constant. But for the long-term not only the address itself but also the address method for digital objects might change.

Therefore it is necessary to apply an unique, location independent identification SYSTEM in conjunction with a method for permanent access to digital objects. This can be achieved using PIs.[2]

The efforts of Die Deutsche Bibliothek concerning the introduction of PIs are motivated by several aspects.

- The rapid rise in the number of online publications and the increasing utilization of paperless distribution channels requires the expansion of the mandate for collecting, indexing and archiving digital resources to ensure long-term access.
- The new possibilities to create, present, describe and publish digital content have been caused by an increasing demand for permanent addressing, citation and unique identification of digital objects.
- The active involvement of Die Deutsche Bibliothek in the CDNL (Conference of Directors of National Libraries) and CENL (Conference of European National Libraries) underscore the need to introduce PIs within an international URN name space "National Bibliography Number" (NBN) of national libraries.

These conditions led to the decision to take an active part in the project CARMEN-AP4 "Persistent Identifiers and Metadata Management in Science" (1999-2002)[3] in order to install a test URN management SYSTEM to gain experiences in respect to URN administration and resolution. The positive results demonstrated the necessity to expand the established URN management and to enhance prototyping into a productive SYSTEM. This will be conducted in the framework of the project "Enhancement of Persistent Identifier Services - Comprehensive Method for unequivocal Resource Identification" (EPI-CUR, 2002-2004).

URN management at Die Deutsche Bibliothek

In Germany there are exceptional challenges with respect to a national URN management because of heterogeneous conditions of certain libraries within a federal

organized library community. These conditions have been targeted by establishing a cooperative URN management for online dissertations.

General concept of the URN management within the international name space "National Bibliography Number" (NBN)

Name space "National Bibliography Number" (NBN)

Die Deutsche Bibliothek is assigning URNs within the name space "National Bibliography Number" (NBN)[4]. NBN is an initiative of both the CDNL and CENL devoted to the introduction and management of PIs within the URN SYSTEM.[5] The goal is to develop the NBN as a distributively managed hierarchical name space. The Library of Congress (LoC) is designated as the global registration agency for the name space "NBN". Each individual national library is to be given a sub name space, in which it can assign and administer NBNs independently. The scope for the sub name space of individual national libraries corresponds to the collective commission of individual national library. NBNs administered at Die Deutsche Bibliothek have the following prescribed structure:

urn:nbn:de[...]

The Uniform Resource Name "urn" identifies the resolution routine, "nbn" stands for the officially registered name space, and "de" is the sub name space for Germany.

General URN management concept

Die Deutsche Bibliothek, in co-operation with German library associations and university libraries, established a distributed URN assignment. The assignment of URNs for online dissertations can be done by all participating university libraries. URNs are registered and resolved centrally at Die Deutsche Bibliothek. To ensure that decentralized assigned URNs are unique the hierarchical structure of NBN-based URNs at regional level has been extended. Following recommendations related to the formal number structure of the NBN based URN at national level have been applied

urn:nbn:de:[designation of library associations]:[official notation of university libraries] -[string][check number]

This model of a cooperative URN management enables participating university libraries

- to assign URNs independently and immediately, to adapt the structure of the last part of
- the URN string to the local conditions and
- to delegate autonomous URN assignment to further associated institutions.

Workflow of URN registrationWorkflow description

After an university library has assigned a URN, it is sent to Die Deutsche Bibliothek together with the metadata

set of the online dissertation. The online publication is archived by the Die Deutsche Bibliothek. The URN will be entered added to the catalogue as an element of the title record. The URN reporting process represents a further expansion of the registration process for online dissertations currently in use.

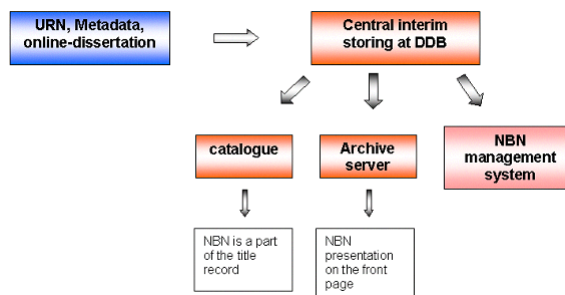


Fig. 1 URN registration workflow at Die Deutsche Bibliothek

German administrative and descriptive Metadata Standard for online dissertation

The information that enables the registration and archiving of online dissertations is based on a metadata set described by the national Metadata Application Profile called "METADISS"[6]. This metadata standard is derived from the international meta data standard "Dublin Core". It determines the representation of the metadata using the HTML markup language.

```

<META NAME="DC.Publisher.CorporateName"
CONTENT="Universit&auml;tsbibliothek Hannover
und Technische Informationsbibliothek">
<META
NAME="DC.Publisher.CorporateName.Address"
CONTENT="Welfengarten 1B, 30167 Hannover">
<META NAME="DDB.Contact.ID" CONTENT="F6000-
0244">
<META NAME="DC.Type" CONTENT="Text.PhDThesis">
<META NAME="DC.Creator.PersonalName"
CONTENT="Schmachtenberg, Oliver">
<META
NAME="DC.Creator.PersonalName.DateOfBirth"
SCHEME="ISO8601" CONTENT="1970-12-12">
<META
NAME="DC.Creator.PersonalName.PlaceOfBirth"
CONTENT="R&uuml;sselsheim">
<META NAME="DC.Title" LANG="eng"
CONTENT="Nitric oxide in the olfactory
epithelium">
<META NAME="DC.Language" SCHEME="ISO639-2"
CONTENT="ger">
<META NAME="DC.Contributor.CorporateName"
CONTENT="Hannover, Universit&auml;t,
Fachbereich Biologie">
<META NAME="DC.Date.Accepted" SCHEME="ISO8601"
CONTENT="2001-06-25">
<META NAME="DC.Subject" SCHEME="DNB-
Sachgruppe" CONTENT="32">
<META NAME="DC.Identifier" SCHEME="URL"
CONTENT="http://edok01.tib.uni-hannover.de/
edoks/e01dh01/332175294.pdf">
<META NAME="DC.Format" SCHEME="IMT"
CONTENT="application/pdf">
<META NAME="DC.Identifier" SCHEME="URN:NBN:DE"
CONTENT="urn:nbn:de:gbv:089-3321752945">
  
```

```
<META NAME="DDB.Contact"
CONTENT="ursula.krys@tib.uni-hannover.de">
<META NAME="DDB.Identifizier" SCHEME="URL"
CONTENT="http://deposit.ddb.de/cgi-bin/
dokserv?idn=962820598&dok_var=d1&dok_ext=pdf&f
ilename=962820598.pdf">
```

Fig. 2 Example of a metadata record

Access and Presentation of URNs within bibliographic information services at Die Deutsche Bibliothek

URNs administered by Die Deutsche Bibliothek are accessible about different channels:

- The national bibliography contains URNs. Furthermore the "Automated Library Exchange Format" (MAB) is used to deliver URNs about several bibliographic services.
- Within the online catalogue a URN is searchable by selecting the research category "Standardnummer" (standard number e.g. ISSN, ISBN and URN). At present URLs are replaced with URNs. Additionally to URNs the reference link to the intermediate front-page of an archived online dissertation at Die Deutsche Bibliothek will be visible.
- URNs are presented in the WebOPAC as a visible string e.g. "urn:nbn:de:gbv:089-3321752945" that refers to its URL presentation in the form of <http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:gbv:089-3321752945>

Titel: Nitric oxide in the olfactory epithelium [Elektronische Ressource] / von Oliver Schmachtenberg, Oliver Schmachtenberg, Oliver
Verfasser: Schmachtenberg, Oliver
Erscheinungsjahr: 2001
Hochschulschrift: Hannover, Univ., Diss., 2001
Persistent Identifier: urn:nbn:de:gbv:089-3321752945
URL: Archivserver Der Deutschen Bibliothek
Sachgruppe: 33 Medizin | 32 Biologie

<http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:gbv:089-3321752945>

Fig. 3 Example of a title record of an online dissertation within the online catalogue

The intermediate front-page contains also a URN. The intention is to inform the user about the existence of a URN for reliable citation.

Archivserver
deposit.ddb.de Die Deutsche Bibliothek

Autor: Schmachtenberg, Oliver
Titel: Nitric oxide in the olfactory epithelium
Dissertation: Hannover, Universität, Fachbereich Biologie, 2001
URN (NBN): urn:nbn:de:gbv:089-3321752945

Dokument im Format: pdf
Dateigröße: ca. 2,0 MB
MDS-Fingerprint des Dokuments

[Dateibetrachter \(Viewer\)](#) zur Präsentation einiger Dokumentformate
[Informationen zum MDS-Fingerprint](#)

Dublin Core Metadaten des Dokuments

Fig. 4: Example of an intermediate front page of the archived online dissertation at Die Deutsche Bibliothek <<http://deposit.ddb.de/cgi-bin/dokserv?idn=962820598>>

Experiences

The general concept of the decentralized URN assignment has proven its worth with respect to the possibility of individualized URN assignment at university libraries. In this context the application of a check number resulted in the discovery of several mistakes of an URN string delivered by an university library. Therefore the check number will continue applying.

In Germany several software SYSTEMs support URN assignment and registration process at Die Deutsche Bibliothek e.g. OPUS (Online Publications of the University of Stuttgart).[7] These condition caused an increasing number of university libraries to participate at the URN management.

Until March 2003 about 12% of URLs registered by university libraries had to be updated. Furthermore, four university libraries announced a complete reorganization of the online publication storage structure. Furthermore, several URLs are temporarily unavailable every 1-2 weeks. During the last 19 months we identified 3 unstable server and three server which were temporarily unavailable. The experience with respect to the availability of URLs suggest the application of the link check procedure every night.

To guarantee permanent document access in case an URL is temporarily unavailable the resolution service refers automatically to the URL of the archived online dissertation at Die Deutsche Bibliothek.

Besides Die Deutsche Bibliothek, at present there are about 15 university libraries which have integrated the application of URNs by means of [http://nbn-resolving.de/urn/resolver.pl?urn=urn\[...\]](http://nbn-resolving.de/urn/resolver.pl?urn=urn[...]) in their local catalogues. This number is steadily increasing.

Current activities at Die Deutsche Bibliothek

The positive response, and acceptance by the university libraries in respect to the application of URNs, and the increasing interest in the implementation of persistent identifiers led to the initiation of the new project "Enhancement of Persistent Identifier Services - Comprehensive Method for unequivocal Resource Identification", (EPICUR, 2002 - 2004). This new project will be conducted at Die Deutsche Bibliothek and is funded by the German Ministry for Research and Technology.

The main objectives of the EPICUR project are the expansion of the existing URN management to other document classes e.g. publisher publications and the migration of the existing service to a productive and enhanced SYSTEM environment. Furthermore a complementary application of URNs and DOIs are considered.

Bibliography

- 1 Trusted Digital Repositories: Attributes and Responsibilities. An RLG-OCLC Report. May 2002, RLG Mountain View, CA, May, 2002 www.rlg.org/longterm/repositories.pdf.
- 2 Further information regarding Persistent Identifiers . www.persistent-identifier.de.
- 3 Project CARMEN-AP4 "Persistent Identifiers and Metadata Management in Science". www.bis.uni-oldenburg.de/carmen_ap4/index_eng.html.
- 4 Juha Hakala: Using National Bibliography Numbers as Uniform Resource Names. October 2001, www.ietf.org/rfc/rfc3188.txt.
- 5 Titia van der Werf : CDNL/CENL activities with identifiers. May 2000, International Federation of Library Associations and Institutions, ifla.org/IV/ifla66/papers/033-82e.htm.
- 6 METADISS - Format des Metadatensatzes für Online-Hochschulschriften Der Deutschen Bibliothek und des Projekts "Dissertationen Online" deposit.ddb.de/metadiss.htm.
- 7 Online Publications: Electronic Papers, Thesis and Dissertations of the University of Stuttgart . elib.uni-stuttgart.de/opus/doku/english/.