

Aachen, 20.11.2001



Fortbildungsveranstaltung: „Persistent Identifier“

Kathrin Schroeder
Die Deutsche Bibliothek



Gliederung

1 **Persistent Identifier**

- Persistent Identifier im theoretischen Kontext
- Beispiele für Persistent-Identifer-Systeme

2 **Persistent-Identifier-Implementation**

- Das Projekt CARMEN-AP4: „Persistent Identifier and Metadata Management in Science“
- Persistent Identifier an Der Deutschen Bibliothek

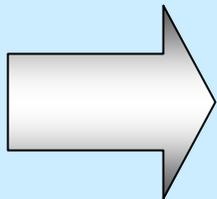
3 **Persistent Identifier im Kontext weiterer Anwendungen**

4 **Exkurs: Metadaten**



Ausgangssituation im Bereich elektronischer Publikationen

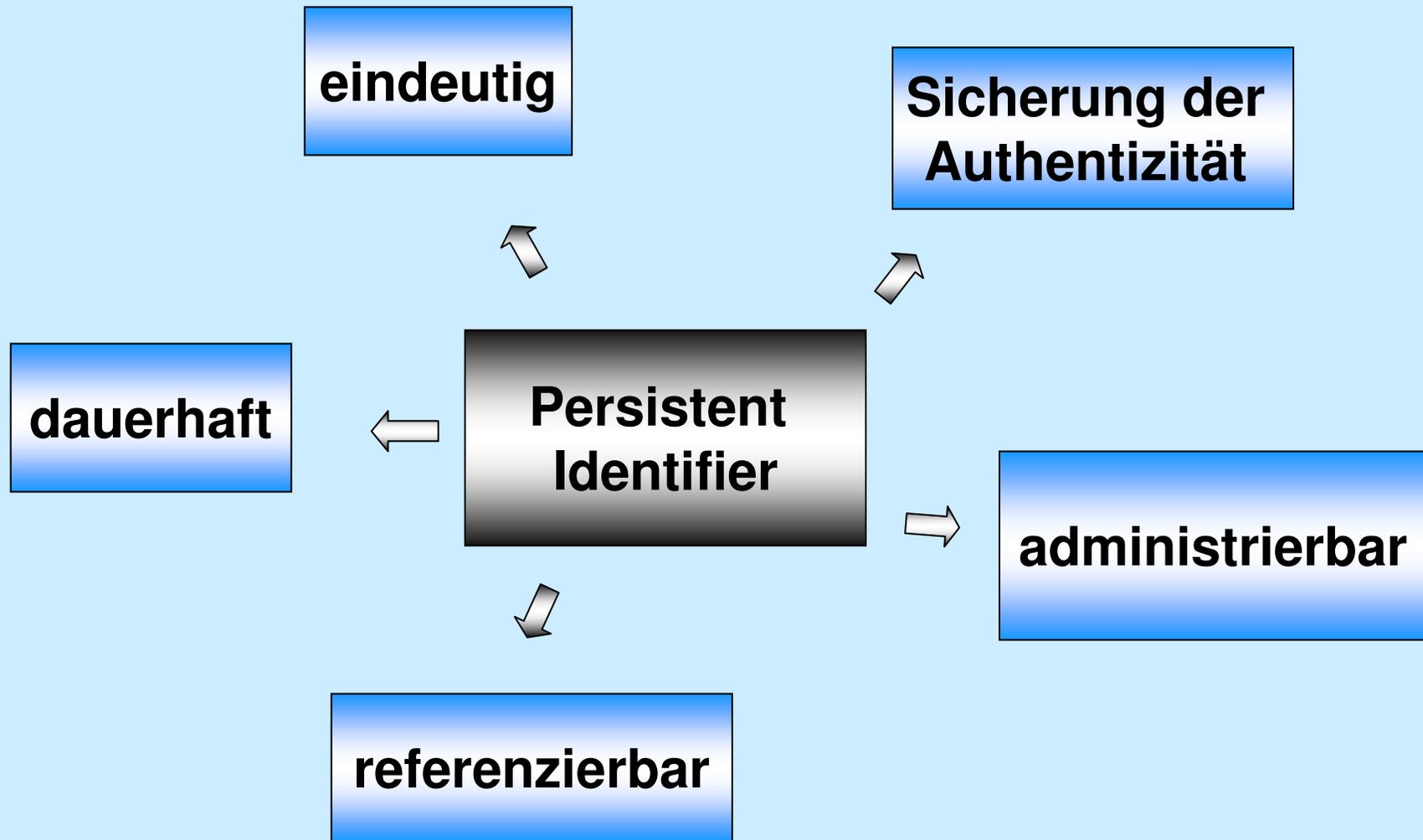
- Steigende Zahl von elektronischen Publikationen
- mangelnder Rechtsschutz
- Probleme mit dem „Uniform Resource Locator“ (URL)
 - Veränderung des Standortes
 - Zitierfähigkeit



Verwendung von **dauerhaften und eindeutigen Adressierungsverfahren** für Online-Publikationen



Lösung: Verwendung von dauerhaften Bezeichnern

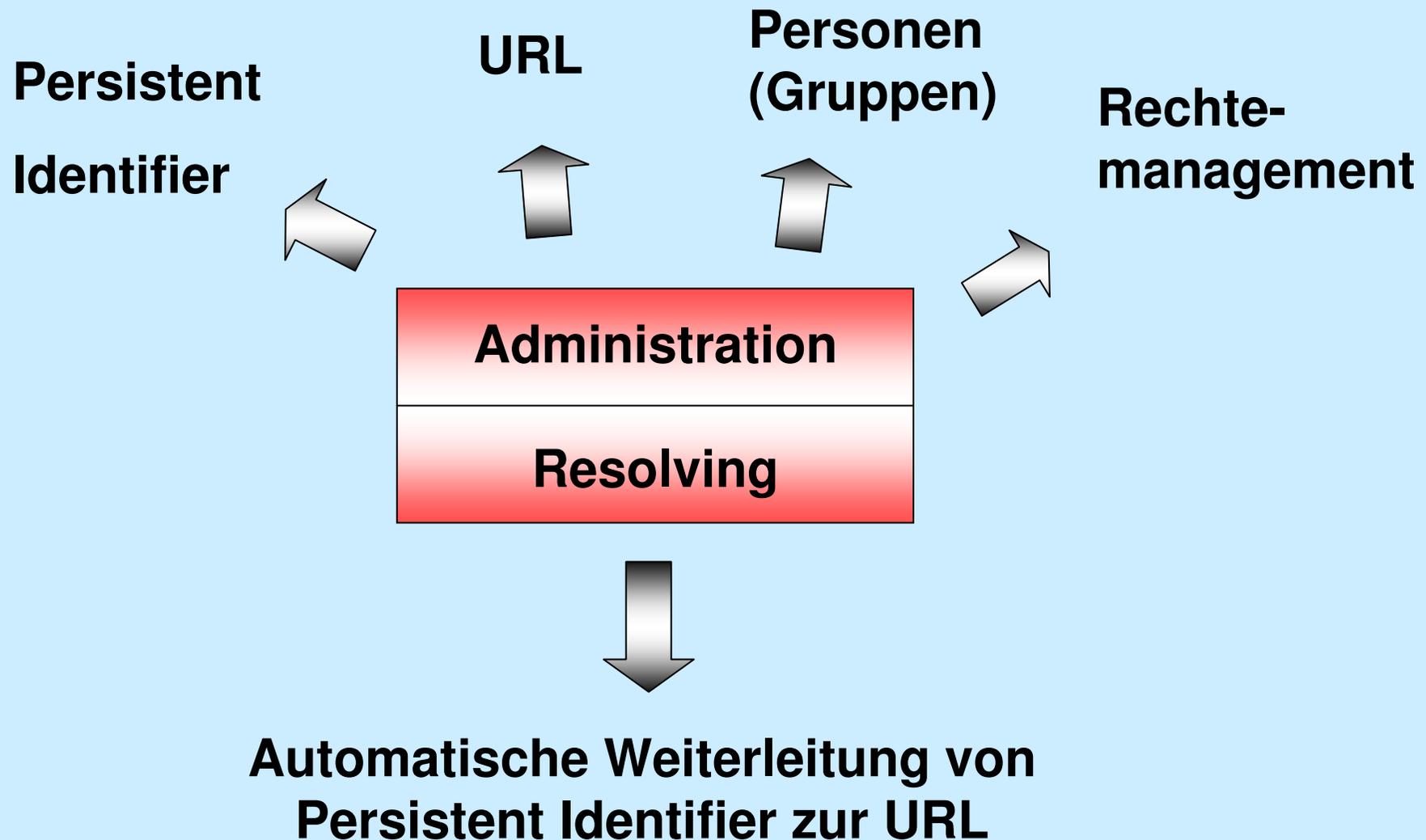


Voraussetzungen für Persistent Identifier Implementierungen

- **Registrierungsinstanz**
- **internationale Konventionen**
 - Persistent-Identifier-Syntax
 - Namensgebung
 - **Zuordnung von Persistent Identifier und URL**
- **Aktualisierung von URLs**
- **Administrierung von Gruppen mit unterschiedlichen Rechten**



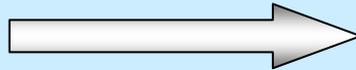
Verwaltungssystem für Persistent Identifier



Resolving-Prozess

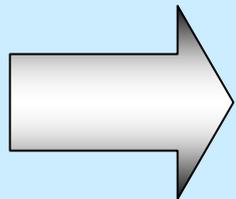
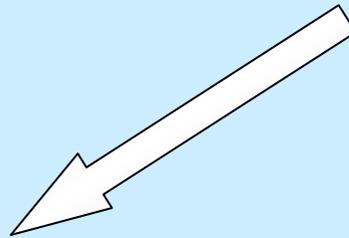
Endnutzer

PI-Anfrage



Verwaltungs-
system

Online-Dokument



Umleitung einer PI-Anfrage an eine **stabile URL**

vorhandene Resolving-Mechanismen

- Installationsaufwand durch den Endnutzer
- **Direkte** PI-Eingabe möglich

Proxy-Server

PlugIn

- **Kein** Installationsaufwand
- Persistent-Identifizier-Eingabe in ff. Syntax:

http://<resolver-adresse>/<PI>

**Persistent-URL
(PURL)**



zukünftiger Resolving-Mechanismus

Direkte PI-Eingabe in die Adresszeile des
Browsers

**internationale
Auflösung**

Gegenwärtiger Resolving-Mechanismus für Praxisanwendungen

http://<resolver-adresse>/<PI>

**Persistent-URL
(PURL)**



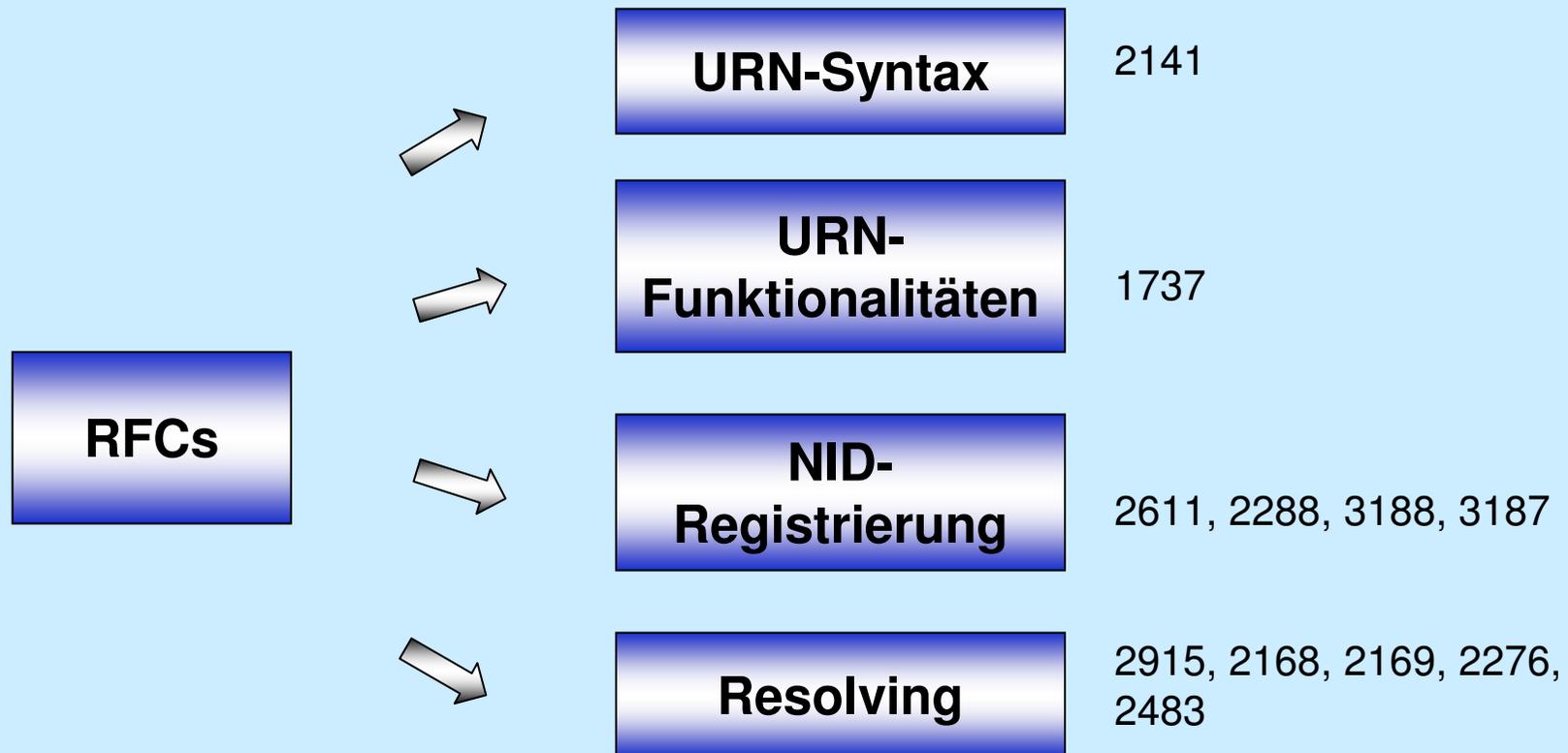
Pause

Beispiele für Persistent-Identifizier-Systeme

- ⇒ Uniform Resource Name (URN)
- ⇒ Digital Object Identifier (DOI)

Der „Uniform Resource Name“ (URN) als Beispiel für ein Persistent-Identifizier-System

Kontrolle durch „**Internet Engineering Task Force**“ (IETF)



Syntax von „Uniform Resource Names“ (URN)

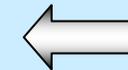
- **allgemeiner** Ansatz
- basiert auf **hierarchischem** Aufbau von Teilbereichen



URN:NID:SNID-NISS

NID: Namensraum (Namespace Identifier)

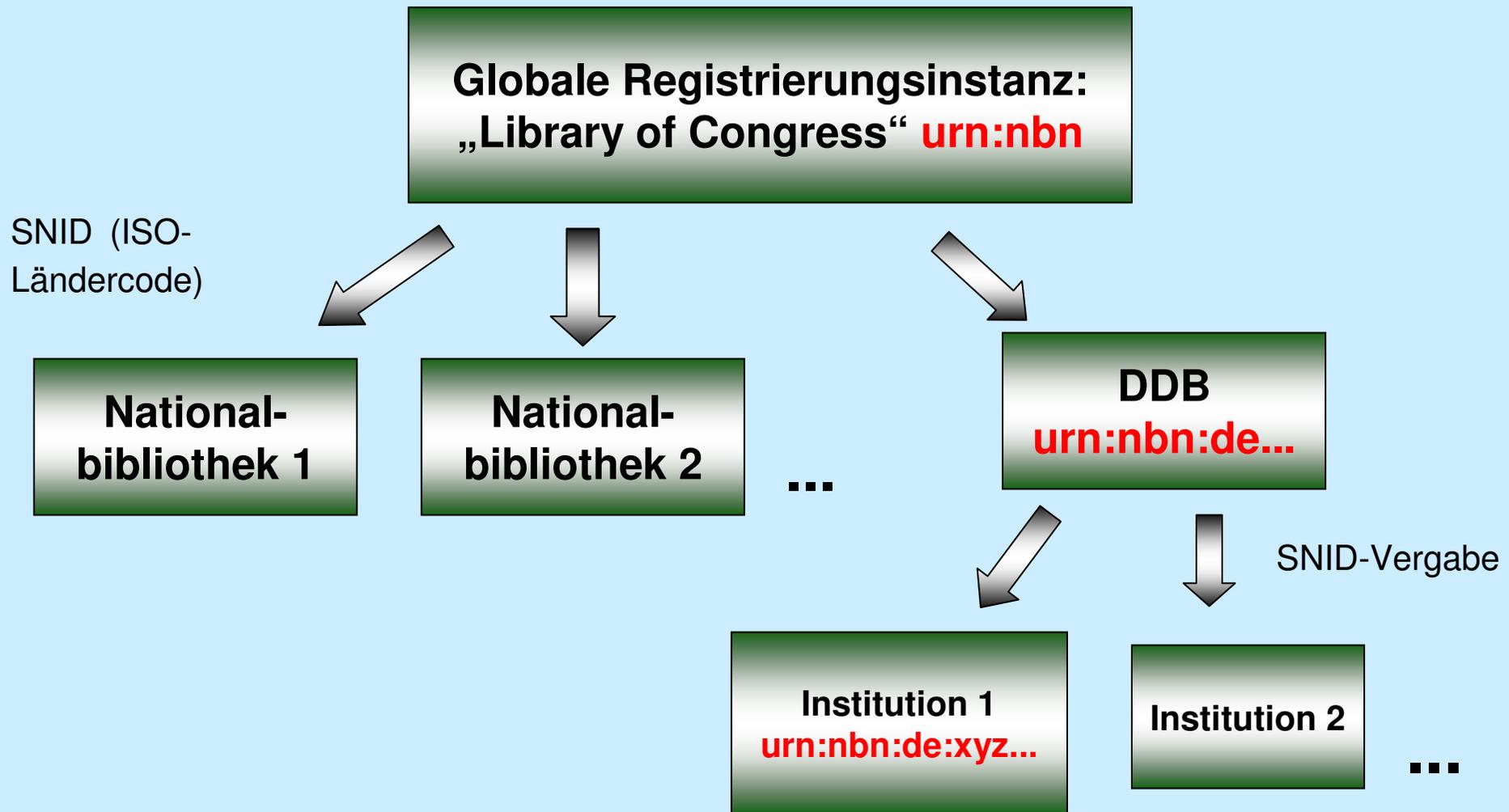
Beispiele: ISBN, ISSN und **NBN**



SNID: Teilbereich aus dem Namespace (Subnamespace oder Sub-Prefix)

NISS: Namespace Specific String

URN-Namespace „National Bibliography Number“ (NBN) URN:NBN...



<ftp://ftp.isi.edu/in-notes/rfc3188.txt>

20.11. 2001

14

Die Deutsche Bibliothek



Registrierte URN-Namespaces

standardisierte **Registrierung** (RFC 2611, 2026)

– formelle Namespaces

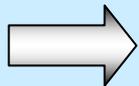
RFCs: IETF, PIN, ISSN, OID, NEWSML, OASIS, XMLORG, publicid, ISBN, NBN

Drafts: SICI, GEO

– informelle Namespaces

z.Z. 3 (Harvard University Library, Rob Akscyn - Knowledge Systems,
Michael Mealling)

–experimentelle Namespaces



<http://www.iana.org/assignments/urn-namespaces>

<http://www.uri.net/urn-nid-status.html>



Digital Object Identifier (DOI) als Beispiel für ein kommerzielles Persistent-Identifizier-System

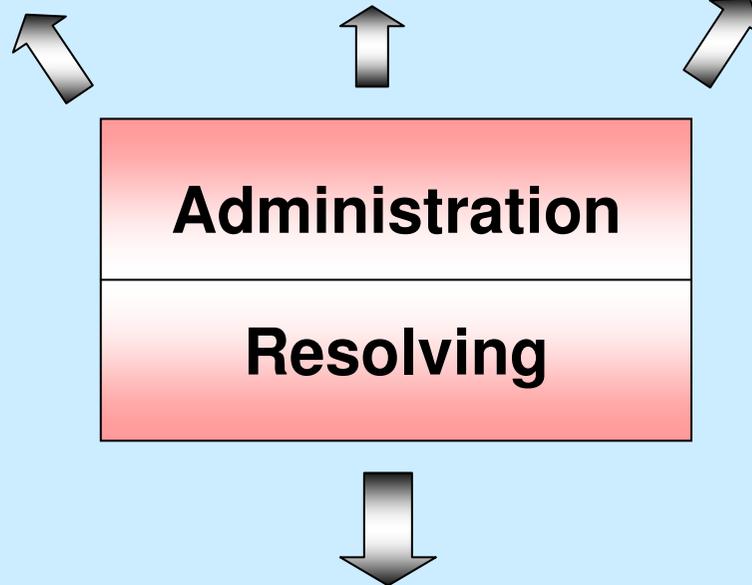
DOI-Foundation (DOI-F) (1996 ff.; <http://www.doi.org/>)

Teilnehmer: ca. **120 Verlage / Organisationen**

Identifizier

Metadaten

Update-Funktionen



CNRI-Handle-System (Proxy, PlugIn)
<http://www.handle.net/resolver/index.html>



Syntax von DOIs

Standardisierung: ANSI/NISO Z 39.84

Bestandteile:

Präfix / **Suffix**

Beispiele:

10.1000/7

10.1045/march99-bunker

DOI-Request:

10.1000/7 **oder**

<http://dx.doi.org/10.1000/7>



Pause

Persistent-Identifizier-Implementation

- ⇒ Das Projekt CARMEN-AP4
- ⇒ Persistent Identifier an
Der Deutschen Bibliothek

Das Projekt CARMEN-AP4: „Persistent Identifier and Metadata Management in Science“

Teilprojekt der SFM **CARMEN** „Content Analysis, Retrieval and Metadata: Effective Networking“

Ziele von CARMEN-AP4:

- **Testlauf** zur dauerhaften und eindeutigen Identifikation von Online-Publikationen durch **Persistent Identifier**
- **Vergabe und Verwaltung** der Persistent Identifier
- Anwendung verschiedener **Resolving-Mechanismen (Proxy, Persistent-URL, NAPTR)**

BIS Oldenburg



20.11. 2001

FIZ Karlsruhe



SUB Göttingen



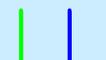
20

UB/TIB Hannover



Die Deutsche Bibliothek

Die Deutsche Bibliothek



CARMEN-AP4 in Der Deutschen Bibliothek

- seit September 2001: Verfahren zur exemplarischen Vergabe und Verwaltung von Persistent Identifiern des **URN-Subnamespaces „NBN:DE“**
- zunächst am Beispiel von **Online-Dissertationen**
- Simulation des **Resolving-Mechanismus Persistent-URL**



NBN-Nummernstruktur im Projekt

Expertentreffen mit Vertretern der **Bibliotheksverbände**
„Gestaltungsmöglichkeiten eines gemeinsamen URN-
Namespaces „National Bibliography Number (**NBN**)“

URN:NBN:DE:[SNID1]:[SNID2]-[NISSP]

SNID1: Kürzel der Bibliotheksverbände (GBV, HEBIS, HBZ, BSZ, KOBV, BSB)

SNID2: Bibliothekssigel der Universitätsbibliotheken (z.B.20)

NISS: - fortlaufende Nummer für Universitätsbibliotheken aus dem BVB, GBV, HEBIS, HBZ, KOBV

- alphanumerischer String für Universitätsbibliotheken des BSZ

P: Prüfziffer



NBN-Darstellungsform im Projekt

URL-Darstellung

- als Hyperlink

urn:nbn:de:[...]



http://<resolver-adresse>/urn:nbn:de[...]
http://nbn-resolving.de/urn/resolver.pl?urn=nbn:de[...]

- ohne Hyperlink (Printmedien etc.)

urn:nbn:de[...]
[http://nbn-resolving.de/urn/resolver.pl?urn=nbn:de[...]]

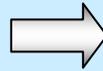


Dokument-Registrierung

Universitätsbibliothek

Die Deutsche Bibliothek

**NBN, Metadaten,
Online-Dissertation**



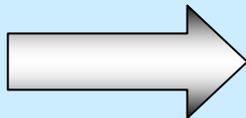
**Zentrale Zwischen-
speicherung bei DDB**



**Katalog
(DDB)**

**Archiv-
server**

**NBN
Verwaltungs-
system**



**Geschäftsgang zur Registrierung von Online-
Hochschulschriften**

Beispiel: NBN-Meldung an DDB

```
<META NAME="DC.Publisher.CorporateName" CONTENT="Universit&uuml;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&uuml;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.Identifier" SCHEME="URL" CONTENT="http://deposit.ddb.de/cgi-  
bin/dokserv?idn=962820598&dok_var=d1&dok_ext=pdf&filename=962820598.pdf">
```



Beispiel: Katalogeintrag mit einer NBN

0500 Oa
0599 01-10-16 : a
0600 rh
1100 2001
1101 Dguu
1500 /1eng
3010 !123256836!
4000 Nitric oxide in the olfactory epithelium [[Elektronische Ressource]] / von Oliver
Schmachtenberg
4083 =A \$
4085 =g urn:nbn:de:gbv:089-3321752945 
4085 =s ca. 2,0 MB =u <http://edok01.tib.uni-hannover.de/edoks/e01dh01/332175294.pdf>
4204 Hannover, Univ., Diss., 2001
4700 fi
4701 ep/Kai
5050 33;32
E01
7001 09-10-01 : x
7800 076937046
7900 10-10-01 08:13:15.695



Beispiel: DDB-Web-OPAC



Titel: Klonierung und funktionelle Charakterisierung des **Aktinreorganisators p150-Spir** [Elektronische Ressource] / vorgelegt von Ines Maria Otto

Verfasser: [Otto, Ines Maria](#)

Erscheinungsjahr: 2001

Hochschulschrift: Würzburg, Univ., Diss., 2001

URL: [Archivserver Der Deutschen Bibliothek
http://erl.bibliothek.uni-wuerzburg.de/DISS/Biologie/2001/X117840/Otto.pdf](http://erl.bibliothek.uni-wuerzburg.de/DISS/Biologie/2001/X117840/Otto.pdf)

Persistent Identifier: [urn:nbn:de:bvb:20-1178402](http://nbn-resolving.org/urn:nbn:de:bvb:20-1178402)

Sachgruppe: 33 Medizin ; 32 Biologie

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



Beispiel: Frontpage des DDB-Archivservers



Archivserver
deposit.ddb.de

Die Deutsche Bibliothek 

Autor : Möhring, Silke

Titel : Selektion und Vermehrung selbstinkompatibler Winterrapslinien
(*Brassica napus* L.) im Hinblick auf die Entwicklung von
Hybridsorten

Dissertation : Hannover, Universität, Fachbereich Gartenbau, 2000

URN (NBN) : urn:nbn:de:gbv:089-3304534323



Dokument im Format: pdf
Dateigröße: ca. 1,0 MB
MD5-Fingerprint des Dokuments

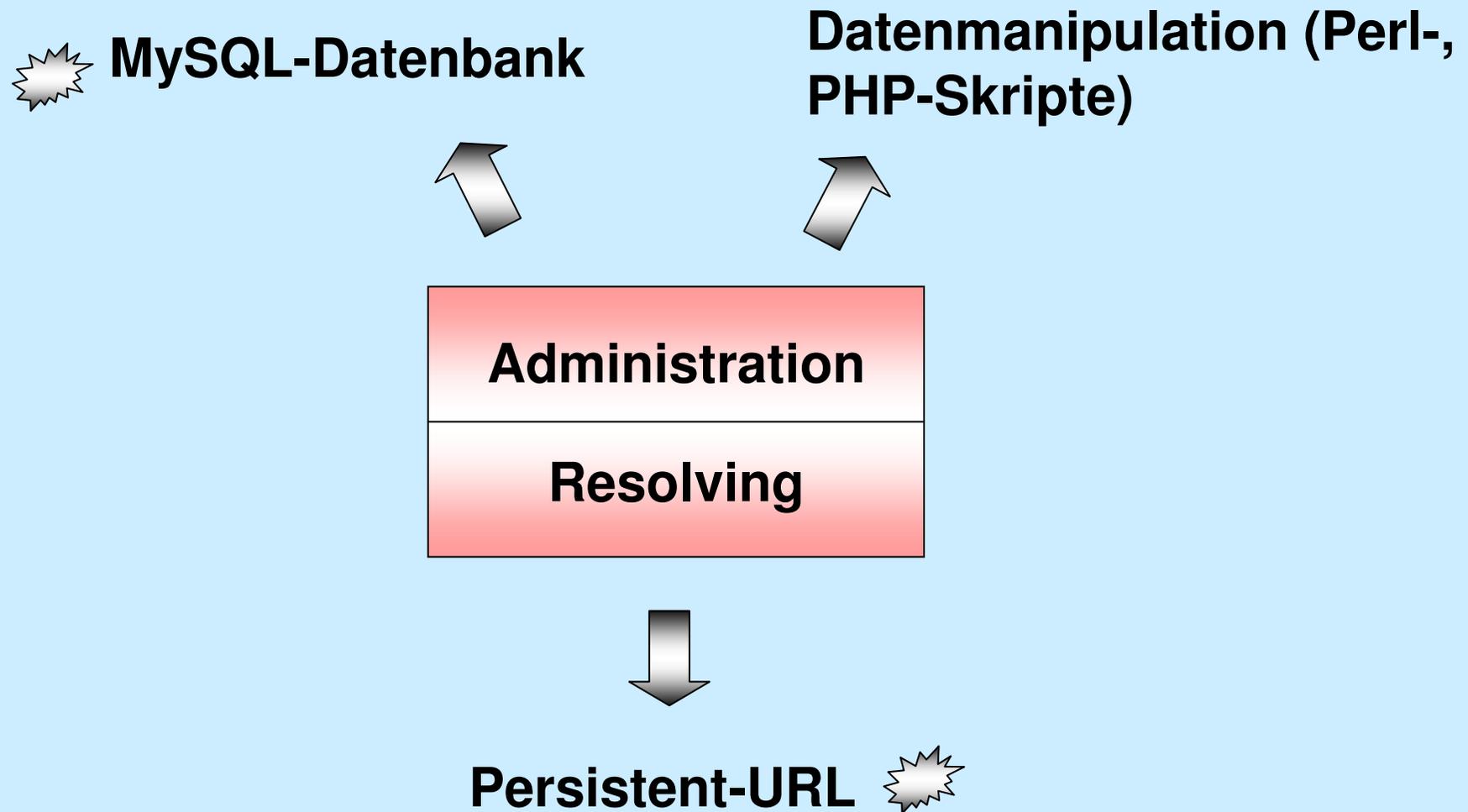


Dateibetrachter (Viewer) zur Präsentation einiger Dokumentformate



Informationen zum MD5-Fingerprint

NBN-Verwaltungssystem



Prüfverfahren im Projekt



URL- Linkcheckverfahren

Periodische
Überprüfung der
Gültigkeit von URLs



Authentizitäts- prüfung von Online- Dokumenten

Periodischer MD5-
Prüfsummenvergleich



Änderungs- und retrospektive Meldungen an Die Deutsche Bibliothek

Universitätsbibliothek

Die Deutsche Bibliothek

URL-Änderungswunsch

**retrospektive Meldung
von NBNs**

Meldung von externen Pls

...



Zentrale Schnittstelle



**Erfassung der Daten
im NBN-Verwaltungs-
system**



Pause

3. Persistent Identifier im Kontext weiterer Anwendungen

Anwendungsmöglichkeiten

- Zitation
- Archivierung
- Effizienter Zugriff bei Mehrfachkopien
- Rechte- / Zugriffsmanagement
- konsistente Links (z.B. in Katalogen, Indizes, Gateways)
- Bookmarks



3. Persistent Identifier im Kontext weiterer Anwendungen

Metadaten

Identifikation, Beschreibung, Nachweis von Objekten

Austausch von dokumentbezogenen Meta-Informationen

Open Archive Initiative (OAI) (<http://www.openarchives.org/>)

OpenURL (<http://demo.exlibrisgroup.com:8888/OpenURL/openurl.html>)

Beispiel

SFX, DOI, CrossRef, OpenURL

(<http://www.dlib.org/dlib/september01/caplan/09caplan.html>) 

4. Exkurs: Metadaten

- ⇒ **Typologie von Metadaten**
- ⇒ **Denkmodell**
- ⇒ **Beispiele**

Ronald Snijder: Metadata standards and information analysis:
a survey of current metadata standards and the underlying models, May 2001
<http://www.geocities.com/ronaldsnijder>



Typologie der Metadaten

Modell von D. Bearman / K. Sochats (1994)

Metadaten

abstrakte Ebene:

- Denkmodell
- Framework

Anwender Ebene:

- resource discovery and retrieval
- terms and conditions
- structure
- content
- use history

Datentransfersyntax



HTML, XHTML,
XML, RDF,
SGML ...

Parallele Verwendung
mehrerer Metadaten-
Schemata mit

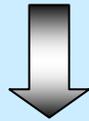
RDF/XML



Kapsel-, Containermodelle für Metadaten

Denkmodell

- **Functional Requirements for Bibliographic Records (FRBR)**



Implementation

Resource Description Framework (RDF)

Framework

- **INDECS Metadata** (<http://www.indecs.org/pdf/framework.pdf>)

(Interoperabilität von Daten im E-Commerce Systemen)



Beispiele: Metadatentransfersyntax und -semantik

- **Online Information eXchange (ONIX)**
- **vCard** (Metadaten für Personen)
- **EAD** (Encoded Archival Description)
entwickelt speziell für das Retrieval von archivierten Dokumenten
- **bibliographische Datenformate:** MAB, MARC21
- **Government Information Locator Service (GILS)**
- **TEI** (Text Encoding Initiative)
internationaler Standard für Erstellung und Austausch von elektronischen Texten im wissenschaftlichen Bereich



Metadaten: Application Profiles (AP)



Metadaten-Elemente aus verschiedenen Namespaces

Beispiele:

Government - AP

Education - AP

Library - AP



Dublin Core Metadata Initiative (DCMI)

Dublin Core (DC) 1995 (<http://dublincore.org/>)

Standardisierung:

- RFC 2413 (DCMI 1998)
- Europäischer Standard: CWA 13874 (2000)
- seit Oktober 2001 ANSI/NISO-Standard Z39.85-2001

Anwendungen in Deutschland:

META-LIB-Projekt (METADISS, Metadaten für Netzpublikationen, METAPERS)

Application Profile (AP)

- AP für Virtuelle Fachbibliotheken (META-LIB-Projekt)

Suchmaschinen / Harvesting

- SWIB
- TheO (Online-Hochschulschriften)
- DC als Grundlage des OAI-Protokolls für Metadaten-Harvesting



Dublin Core

Spezifizierung



15 Elemente:

Title Source
Creator Language
Subject Relation
Description Coverage
Publisher Rights
Contributor
Date
Type
Format
Identifier

Qualifiers

Date:
Alternative
Description:
Table of Contents, Abstract
Date:
Created, Valid, Available, Issued,
Modified
Format:
Extent, Medium
Relation:
Is Version Of, Has Version, Is
Replaced By ...
Coverage:
Spatial, Temporal



DC-Anwendung: METADISS

- DC-Metadatenformat für **Online-Hochschulschriften**
- **Standard** in DDB-Metadaten-Schnittstelle vorgegeben
- Abgestimmt mit „**Dissertationen Online**“ und **Bibliotheksverbänden**
- Kodierung in **HTML 4.0**
- Obligatorische / zusätzliche Angaben
- Angaben sind teilweise **wiederholbar**

Technische Umsetzung von DC / DC-Tools

Kodierung von Dublin Core in RDF/XML

simple Dublin Core von Eric Miller [u.a.]
(<http://dublincore.org/documents/2001/09/20/dcmes-xml/>)

qualified Dublin Core von Stefan Kokkelink [u.a.]
(<http://dublincore.org/documents/2001/08/29/dcq-rdf-xml/>)

Tools zum automatischen Erzeugen von DC-Metadaten

MKDoc und **MMMfT**

Generierung von Metadaten nach qualified Dublin Core, RDF/XML und HTML

DC-Dot

Projekt UKOLN (<http://www.ukoln.ac.uk/metadata/dcdot/>)



Weitere Metadaten-Projekte

Technische Metadaten

- NEDLIB (EU-Projekt, abgeschlossen)

Metadaten zu „Terms and conditions“ und zur Archivierung

- CARMEN AP 2/5

Registrierung der einzelnen Metadaten-Anwendungen

- SCHEMAS (EU-Projekt)
- MMI-DC Metadata Observatory



Perspektiven

Konsistente Linkstrukturen



Generierung von Mehrwert

- Verknüpfung von Datenbanken
- kontext-sensitive Linkstrukturen
- Erstellen von semantischen Beziehungen

Zitierfähigkeit

Weiterführende Informationen

Ansprechpartner in DDB:

Persistent Identifier

Christel Hengel (hengel@dbf.ddb.de, Tel: 069-1525-1401)

Kathrin Schroeder (schroeder@dbf.ddb.de, Tel: 069-1525-1480)

Metadaten

Berthold Weiss (weissb@dbf.ddb.de, Tel: 069-1525-1404)

Links:

Persistent Identifier

http://www.bis.uni-oldenburg.de/carmen_ap4/

<http://www.nla.gov.au/initiatives/persistence/PIcontents.html>

Metadaten

<http://www.geocities.com/ronaldsnijder>

