

# Digital Preservation

## Unique Identifiers

Hannes Kulovits  
Institut für Softwaretechnik und Interaktive Systeme  
TU Wien

<http://www.ifs.tuwien.ac.at/dp>

- Electronic Publications on the Internet
- Advantages:
  - Instant Access
  - Easy and low-cost distribution
- Fundamental Requirement:
  - Unambiguously identify and locate documents
- Problem:
  - Links break (‘404 Not Found’ error)
- Reliable referencing is needed!

- Traditional Mechanisms
  - International Standard Serial Number (ISSN)
  - International Standard Book Number (ISBN)
- Locator Based Mechanisms
  - Uniform Resource Identifier (URI)
  - Uniform Resource Locator (URL)
  - Uniform Resource Name (URN)
  - National Bibliographic Numbers (NBNs)
- Delegating Methods
  - The Handle System
  - Digital Object Identifier
  - Persistent URL (PURL)
  - Archival Resource Key (ARK)

- International Standard Serial Number
- Unique eight-digit number
- Identifies periodical publications (e.g. Magazines)
- Can be encoded as a URN

- International Standard Book Number
- Unique commercial book identifier barcode
- Assigned to each edition and variation of a book
- Consists of 13 digits (since Jan. 1st, 2007), before: 10 digits
  - Prefix 978 or 979 (for 13 digit ISBN)
  - Group identifier code (GIC)
  - Publisher code
  - Item number
  - Checksum character or check digit
- Example: 1-84356-028-3
- Assigned by International ISBN Agency

- Uniform Resource Locator
- Particular form of URIs
- Addressing documents
- Can only be used to locate resources
- Depend on DNS information
- URLs are not persistent
- Problems
  - Links may break
  - No mechanism to handle broken links

- Uniform Resource Name
- Combination of
  - Namespace identifier (NID)
  - Namespace specific string
- Naming scheme for URNs:
  - `,urn: ` <NID> ,: ` <NSS>`
- Example: `urn:isbn:0451450523`
- Internet information architecture
  - URNs
  - URCs (Uniform Resource Characteristics)
  - URLs

- Main functions of a URN
  - Global scope of names
  - Global uniqueness
  - Persistence
  - Scalability
  - Legacy support
  - Extensibility
  - Independence
  - Resolution
  
- Persistence of identifiers
  
- Possibility to resolve them



- National Bibliographic Numbers
- Generally used by national libraries
- Based on URN
- Global registry institution: Library of Congress
- Projects using NBN: DIVA, EPICUR
- Syntax
  - URN:NBN:<ISO country code>-<assigned NBN string>
  - URN:NBN:<ISO country code:sub-namespace>-<assigned NBN string>
  - URN:NBN:<none-ISO prefix>-<assigned NBN string>
- Example
  - urn:nbn:hu-3006

- Distributed persistent naming system
- Conforms to URN framework
- Used by DOI (Digital Object Identifier) system
- Persistent identifier consists of two parts:
  - Naming authority
  - Name (must be unique string to the authority)
- Digital objects on the Internet can be assigned, managed, and resolved by handles.
- Resolved by global handle service

- **Main points**
  - Handles are unique and persistent
  - Handle system supports internationalisation
  - Operations on handle system have to be authorised
- **Syntax:**
  - `<Handle Naming Authority> , / \ <Handle Local Name>`
- **Example:**
  - `152.86/feb2007-rk328642133`
- **Available Services:**
  - `http://hdl.handle.net`

- Allocation of unique digital identifiers to commercial digital publications
- Managed and controlled by Internat. DOI Foundation (IDF)
- DOI names provide
  - Current information
  - Location where they can be found
- Location may change but DOI does not!
- Main functionalities
  - DOIs are persistent
  - DOIs are actionable

- **Syntax**
  - `<DOI Identification> , . ` <prefix> , / ` <suffix>`
- **Example:**
  - 10.1043/182
- **10.1043: prefix, or publisher ID**
  - 10: identifies it as a DOI
  - 1043: identifies the registrant
- **Services**
  - <http://dx.doi.org>

- Persistent URL
- Based on URN specification
- Naming and resolution service for general Internet resources
- Looks like a URL but points to a resolution service
- A PURL is a link to the current location.
- When location of the resource changes, it only has to be updated in the PURL resolver service.
- Main design goals:
  - Promote the use of standard services and protocols
  - Persistence
  - Locators are separated from Internet resources

- **Syntax**

- <protocol><resolver address><name>

- **Example**

- <http://purl.oclc.org/NET/DIGITALISIERUNG.PPN>

- **PURL consists of**

- Protocol
  - Resolver address
  - Name

- **Available services**

- PURL Software downloadable:  
<http://www.oclc.org/research/projects/purl/download.htm>

---

## PURL Information Display

---

**PURL** <http://purl.oclc.org/emls/emlshome.html>  
**URL** <http://www.shu.ac.uk/emls/emlshome.html>

**Maintainers** EMLS

**Creation Date** Thu Jun 20 13:27:08 1996  
**Last Modified** Wed Jul 28 12:29:35 1999  
**Change Category** User\_Edited

---

### History

**Modified by** EMLS  
**Modified** Tue Sep 23 20:08:44 1997  
**Change Category** User\_Edited  
**Reason for Change** server move  
**URL** <http://www.humanities.ualberta.ca/emls/emlshome.html>

**Modified by** EMLS  
**Modified** Thu Jun 20 13:27:08 1996  
**Change Category** User\_Entered  
**Reason for Change** new server  
**URL** <http://unixq.ubc.ca:7001/0/e-sources/emls/emlshome.html>



- Archival Resource Key
- ARKs can identify objects of any type
- Also physical objects such as books
- Designed to allow integration of other schemes and being integrated into other schemes
- Functionalities
  - Link from object to metadata,
  - to object itself, and
  - to a maintenance commitment from the current server

- [ `http://`<NMAH> `/ ` ] `ark:/`<NAAN> `/ `<Name> [<Qualifier>]
- **Example**
  - <http://foobar.zaf.org/ark:/13030/654x321/s3/f8.05v.tiff>
- Name Assigning Authority (NAA)
- Name Assigning Authority Number (NAAN)
- Name Mapping Authority Hostports (NMAH)
  
- NAANs have 5 digits
- 13030 ... California Digital Library

- ISBN, ISSN no resolution mechanism
- URI and URLs not persistent; links may change or break
- URNs not directly available
- NBN is a namespace, exclusively assigned by national libs.
- DOIs, Handles, PURLs, ARKs rely on resolving mechanisms