Persistent Identifiers (PID) @NLB

Ganesh Yanamandra
National Library Board, Singapore

What is a PID

- A PID provides a means of uniquely and persistently identifying a digital object and associating it with related data in a structured and extensible way

- The Persistent Identification System (PID) at NLB is based on the guidelines defined by the International DOI Foundation (IDF) in the *DOI Handbook* - http://www.doi.org/hb.html

- NLB's PID system is developed on the Handle System from CNRI as the core underlying technology.

Scenario 1 – current content identification

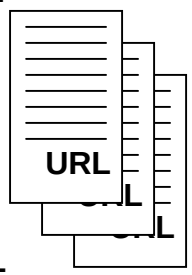
URL

URL

URL

URL

URL



URL

URL

URL

URL

URL

URL

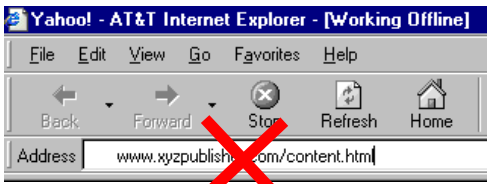
Content

Illustration taken from David Sidman's presentation at National online 2001

Scenario 2 – change in URL

URL

URL



URL

URL URL

URL

URL

URL

404
File not found

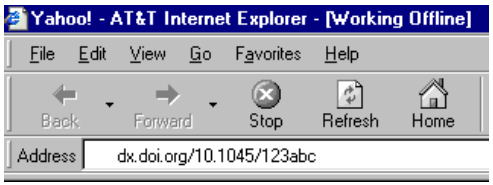
Content

Illustration taken from David Sidman's presentation at National online 2001

Scenario 3 – Persistent Identification

PID

PID



PID

PID PID

PID

PID

PID

PID

PID

PID PID

Resource Provider

PID directory

Content

Illustration taken from David Sidman's presentation at National online 2001

What is a PID meant to do

More than an identifier…

Identify

IDF states that the DOI syntax can include **any existing identifier**, formal or informal, of any entity eg

10.2341/0-7645-4889-1
10.5678/978-0-7645-4889-4
10.1000/ISBN 0764548891
10.1234/Godfrey_presentation
10.2224/2003-1-29-CENDI-DOI

Describe

PID metadata can be **of any type, standard or proprietary**
eg
OnixForBooks
OnixForSerials
IEEE/LOM
MARC
Dublin Core
Proprietary scheme

(but if you want to interoperate with anyone else in the DOI network, you map to the **<indecs> Data Dictionary (iDD)**.

Resolve

The Handle resolution technology allows you to access **any kind of Service** associated with your PID.
eg

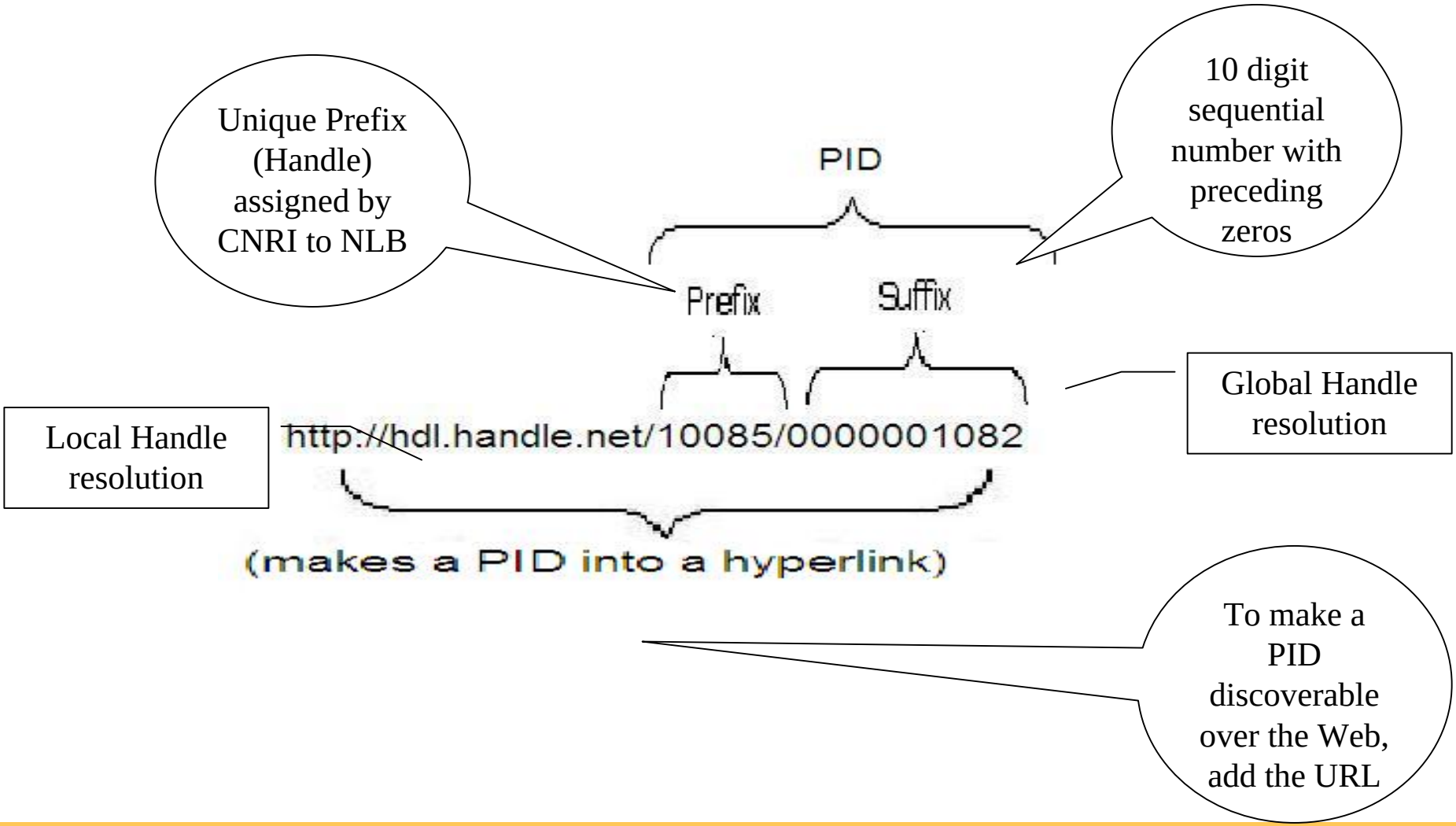
- Retrieve ebook
- Articles and reviews
- Books from same author
- Accompanying music
- Accompanying images
- Accompanying moving images
- Metadata
- Rights information
- Email query to publisher

A package of services is an Application Profile

NB: these Services themselves include metadata services

Illustration based on presentation by Norman Paskin, 2004

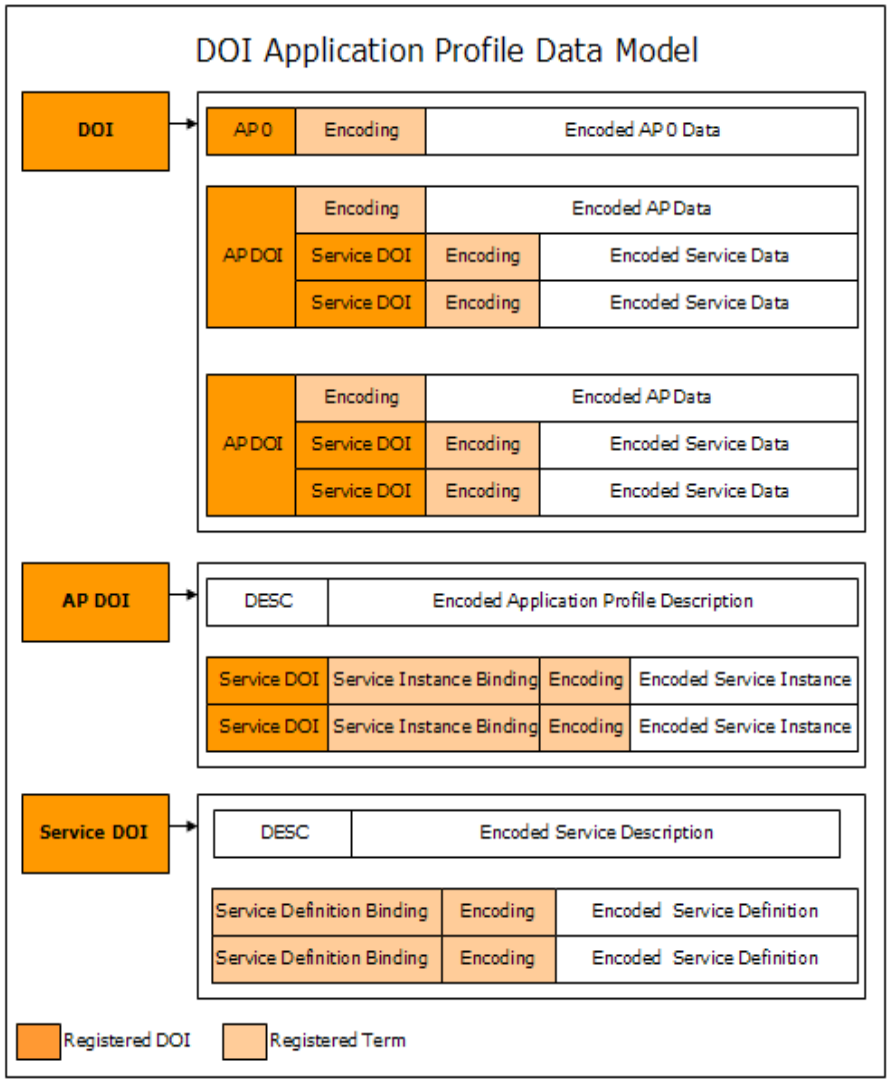
Structure of a PID



How is the Persistent Identifier system (PID) currently implemented in NLB

- PID system implementation at NLB is based on the Handle system from CNRI
- Data model, XML schema, Services are based on IDF specifications
- Kernel Metadata elements are mapped using NLB's Application profile and <Indecs>
- PID system has web services as well as conforms to standard http resolution
- Resources can be resolved using Global Handle System as well as NLB's PID server.
- ODRL Policy's are assigned an PID number and embedded within the metadata <dc:rights> elements to enhance DRM enforcement on Digital resources

NLB's PID data model based on IDF framework



PID

Zero AP (0)

Resolves to a URN (URL/URI)

Base AP (1)

Resolves to a URN (URL/URI)

GetKMD – Returns the PID Kernel Metadata

Full AP (2)

Resolves to a URN (URL/URI)

GetKMD – Returns the PID Kernel Metadata

GetRights – Returns the Rights statement

More services to be defined…….

Restricted AP (3)

Services to be defined……..

Application Profile for Persistent Identifiers (PID)

Sr#	Element	Description	Mandatory	Datatype	Scheme	Example
1	Identifier	A unique label which makes the Resource referable.Uncontrolled alphanumeric string. Structure may be governed by rules	Yes	String	DC:Identifier	PID:2889 DOCID241646124 SIL842374 ISBN 0385517238
2	resourceName	A non-unique label which makes the Resource referable.	Yes	String	DC:Title Has attributes which relate to DC:Title DC:Language	“Singapore: Community Living”
3	PrincipalAgent	An Entity acting in relation to the Resource. Identifier or Name of a related Agent.	No	String	DC:Creator DC:Contributor DC:Publisher	
4	AgentRole	Role of the Agent	No	String	NLB Terms	Note: This is an Open list
5	StructuralType	A classification or type to which the Resource belongs, or a quality which it has.Allowed value Code or Name.	Yes	String	IDD Terms	Note:This is a restricted list, no other values can be used. Abstraction Performance Digital Physical Restricted
6	mode	The principal sensory mode(s) in which a resource is intended to be perceived.	Yes	String	IDD Terms	Note:This is a restricted list, no other values can be used - Abstract Audio Visual AudioVisual Tangible Restricted
7	ResourceType	The general type of a resource.	Yes	String	DC:Type	DCMI Type List
8	Registrant	The identifier assigned to the entity, which is requesting the DOI registration.	Yes	String	NLB Terms	Note: This ID can again be a DOI Identifier, which is the result of a successful registration by the content provider.
9	ApplicationProfile	Application Profile to which this DOI should be assigned to.	Yes	String	NLB Term	Note : These are the Application Profiles that would be created in the NLB’s DOI System implementation. 0 - Zero 1 - Base 2- Full 3- Restricted
10	Rights	Describes the Rights for this resource.It can either be: a)Identifier to a copyright statement in XRML/CopyRightMD if the Application Profile is Restricted b)A text statement.	Yes	String		Note - we are not implementing option a in phase 1

How can you register a PID – checklist for resource providers

a) Register to the PID system as a Resource Provider – mandatory for identification

b) Verify the Primary Contact & his email address – mandatory for communication

c) Note Down the registrant Code returned in the response – mandatory for registration

d) Utilize the Registrant code when you are building interface to the PID system.

e) Conduct a mapping using the Table against your own metadata elements.

f) Decide how will you translate the PID restricted values against your own metadata values.

g) Review the WSDL or Web service definition file from the PID system team.

h) Prepare to participate in the Integration testing on development environment with some test data for both single & batch registration.

Where we are now ?

- PID system is implemented enterprise wide
- NLB has obtained the unique prefix from CNRI –10085
- What does it mean – Any application can use the PID services
- What are the available PID services –
  - Registration request for a new, unique PID number
  - Update details for an registered PID
  - Batch Register
  - Resolution of a PID
    - Seek resolution of the rights statement associated with an PID
    - Seek the Kernel Metadata for an registered PID
    - Resolve to the Digital Content directly

PID usage with samples

1) Submit a request for registering an object with valid details in the xml format to the PID system

XML Document

2) Interpret the response and add the PID number to the metadata.

XML Document

3) PID for a Rights policy statement resolved via Global Handle server - http://hdl.handle.net/10085/0000001082

NLB

What is the plan ahead ?

- Build more services and PID AP's
- Address URN resolution
- Enhance URL resolution with verbs

eg.: http://hdl.handle.net/10085/0123456789?get=kmd

http://hdl.handle.net/10085/0123456789?get=rights

….etc.

- Propose the Application Profile for Persistent Identifiers to DCMI

NLB

Questions ?

email : ganesh_yanamandra@nlb.gov.sg

Thank You !