Overview of PID Systems for Digital Objects: Introduction

August 2016

Views About PID Systems:
Training Course

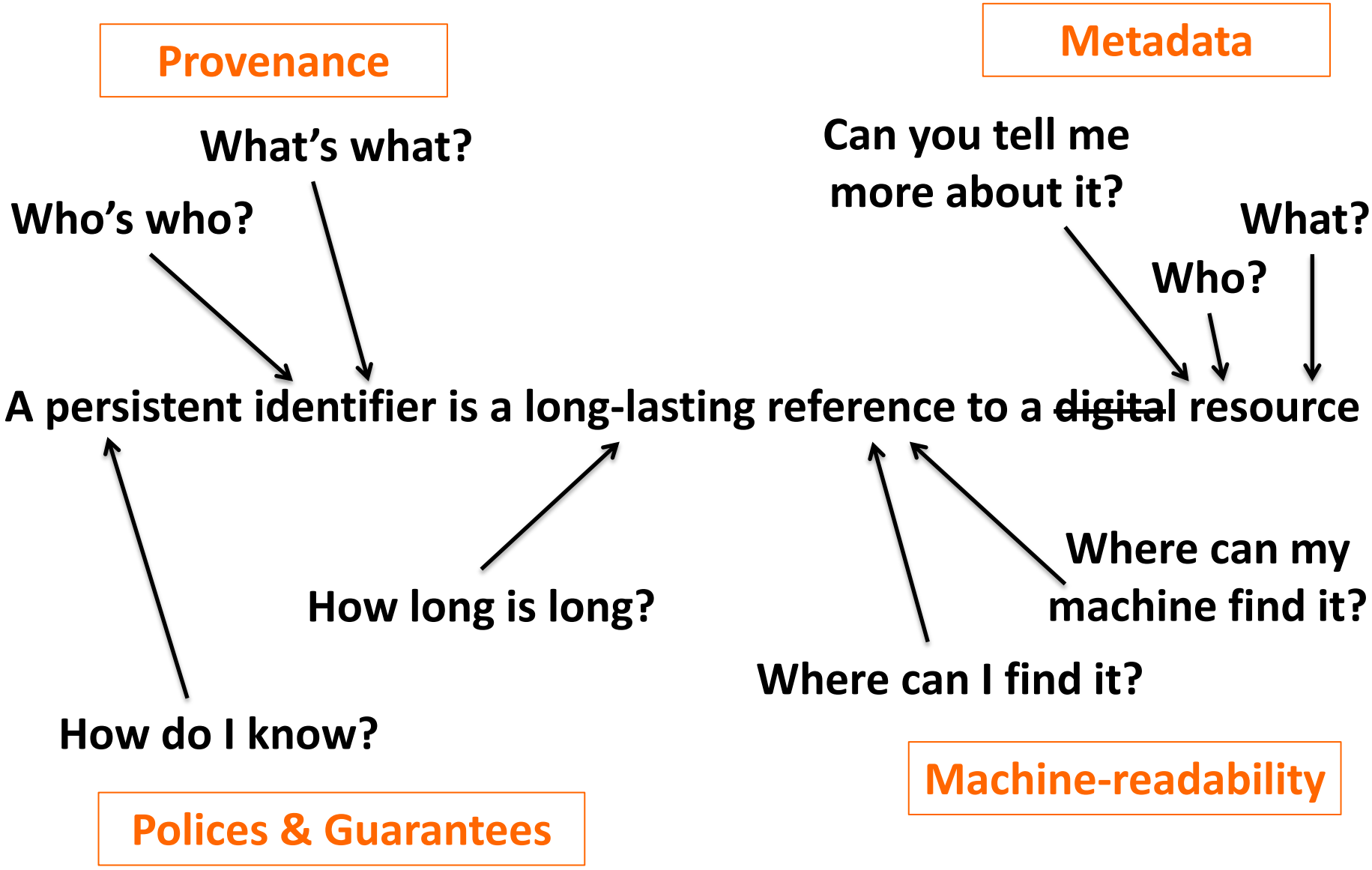
Jonathan Clark
International DOI Foundation (IDF)
j.clark@doi.org
www.doi.org

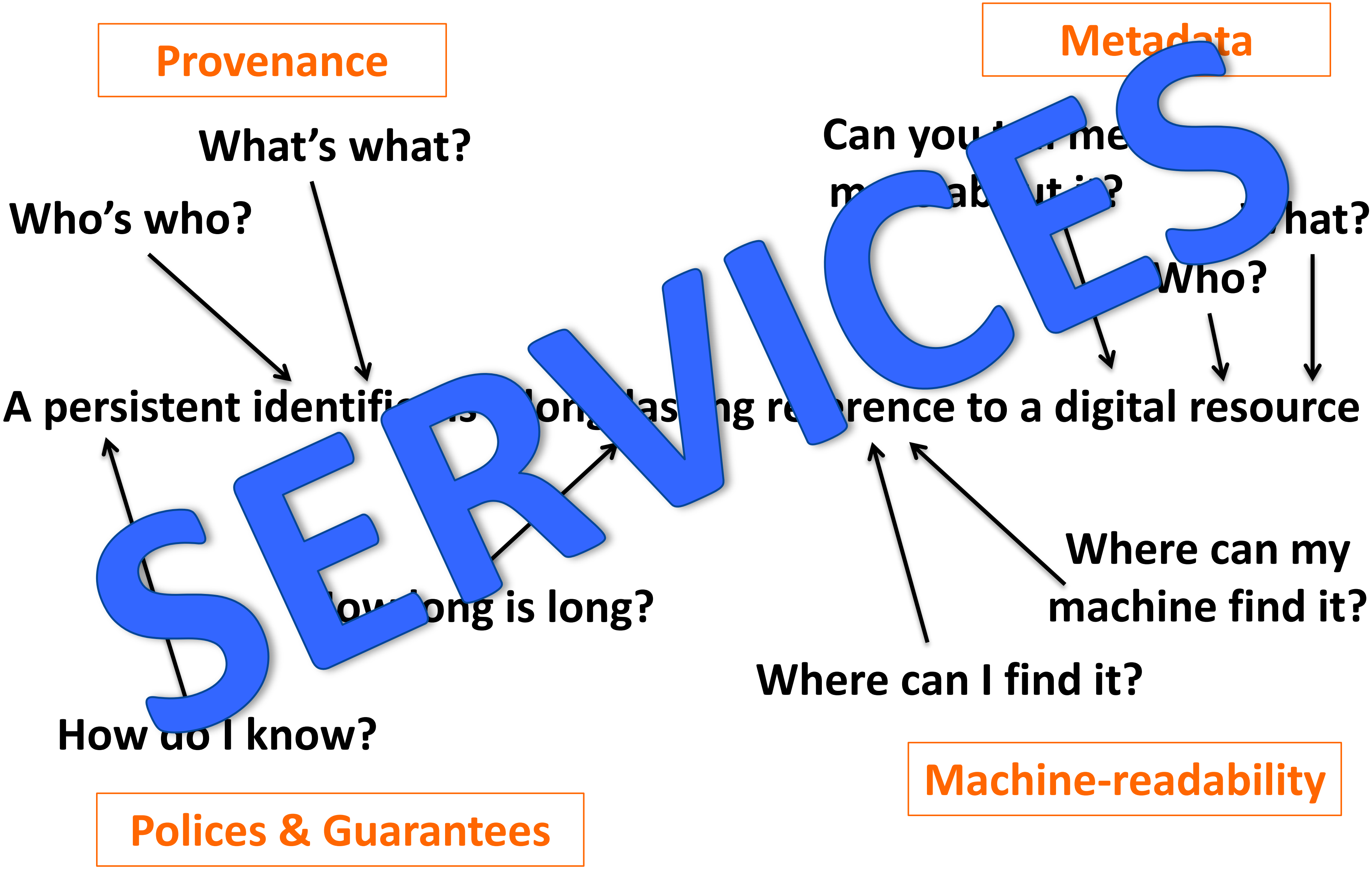

It says:
"Error 404
Page not found"

What's what?

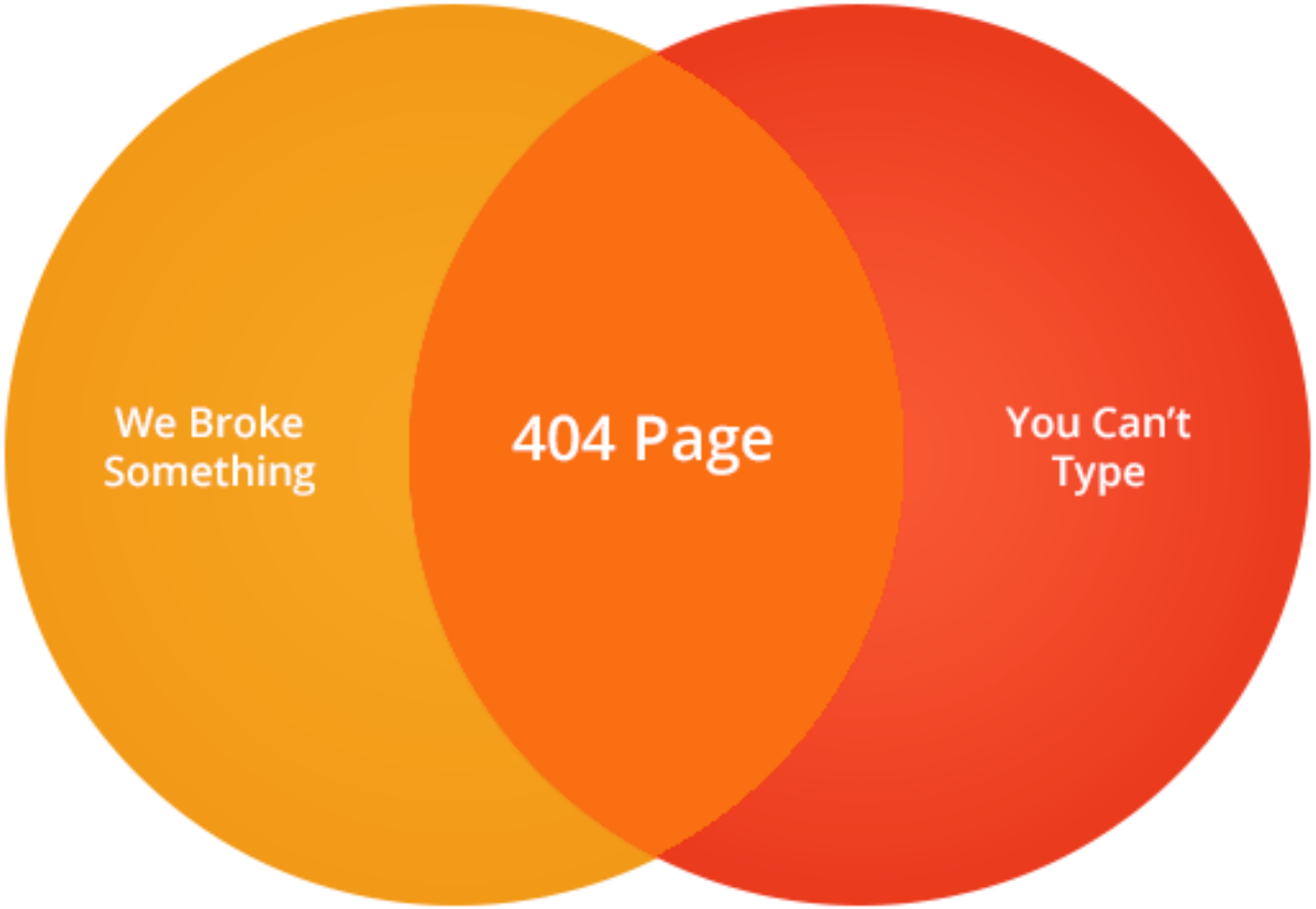
A persistent identifier is a long-lasting reference to a digital resource

Where can I find it?


Provenance
Metadata
What's what?
Who's who?
Can you tell me more about it?
What?
Who?
A persistent identifier is a long-lasting reference to a ~~digital~~ resource
How long is long?
Where can my machine find it?
Where can I find it?
How do I know?
Polices & Guarantees
Machine-readability



Provenance
Metadata
What's what?
Who's who?
Where can my machine find it?
Where can I find it?
Machine-readability
Polices & Guarantees
SERVICES



A human problem needs a human solution

A trustworthy PID system must

- be maintained by a dedicated and reliable team,
- be based on a transparent sustainable business model,
- be provided by a non-profit organisation,
- be subject of regular quality assessments by external parties,
- be governed by international boards,
- be based on open standards,
- be based on a redundant and secure architecture,
- support a huge address space (comparable or even larger than IPv6) and
- support an openly documented API optimally supporting accepted data models.

Handles and Handle System

Established in 1994
Non-commercial decentralized identifier resolution system
Governed and managed by DONA Foundation
DONA is not-for-profit member organisation
Members are Multi-Primary Administrators (MPA)
Handles are unique and persistent identifiers for Internet resources
Handles are of the form <prefix>/<suffix>
Handle System consists of a repository, a resolution system and a registry
Documented in informational RFCs

DOIs and DOI System

Established in 1996
DOI = Digital Object Identifier
Centralised governance and shared infrastructure
An implementation of the Handle System using reserved 10. prefixes
Governed and managed by International DOI Foundation (IDF)
IDF is not-for-profit member organisation
DOI deposit through Registration Agencies (RAs)
RAs are independent bodies offering services to assigners using DOIs
All RAs must sign RA agreement on the use of IDF System (policies & guarantees)
DOIs are actionable, interoperable, persistent links
DOIs are of the form <10.xxxx>/<suffix> resolvable using https://dx.doi.org/
International Standard: ISO 26324, 1 May 2012

ARKs

University of California
CDL
California Digital Library

Introduced in 2001
ARK = Archival Resource Key
Decentralized identifier system
Developed by California Digital Library (CDL)
Organisations can sign up to become Name Assigning Authority Numbers (NAANs) and run their own resolution infrastructure for ARKs
A complete NAAN registry is maintained by the CDL and replicated at the Bibliothèque Nationale de France and the US National Library of Medicine.
ARK is a URL that provides a multi-purpose identifier given to information objects of any type
ARKs are of the form http://NMAH/ark:/NAAN/Name
NMAH: Name Mapping Authority Host – the organization that currently provides service for the object
ARK provides 3 generic services (Access, Policy and Description) Not a formal standard, all ARKs follow the same structure and workflows

URNs

Introduced in 1994, formalised in 1997
URN = Universal Resource Name
Persistent, location-independent, resource identifiers
No central governance for URN and no central resolving infrastructure
Major national libraries in Europe have established their own subgroup of URN (URN:NBN) and operate a joint resolving infrastructure
Syntax is <URN> ::= "urn:" <NID> ":" <NSS> where <NID> is the Namespace Identifier, and <NSS> is the Namespace
For example: urn:isbn:9789521061547 or urn:ietf:rfc:2648
Work continues on syntax (latest update June 2016)
Documented in informational RFCs (RFC 2141)

URN:NBNs

Introduced in 2001
NBN = National Bibliography Number
Unique and permanent identifiers of digital objects, independent of their storage location
Publication identifier systems used by national libraries in countries such as Germany, Italy, Finland, Norway, and Sweden.
The US Library of Congress is the global registration agency for the NBN namespace
Each national library uses its own NBN strings; there is no global authority which controls them. Thus, NBNs are unique only on national level
Resolvers also operate on a national level
Example: urn:nbn:de:bvb:19-146642
URN namespace for NBNs has been assigned and is described in IETF RFC 3188.

PURLs

Introduced in 1995
PURL = Persistent Uniform Resource Locator
Intended as an interim system to be used until the URN framework is well established
Developed and implemented by OCLC
A PURL looks just like a URL, except it points to a resolution service instead of the actual location of the digital resource. The resolution service then redirects the user to the appropriate URL

URIs

Introduced in ?
URI = Uniform Resource Identifier
An extensible means for identifying a resource within the Web.
Each URI begins with a scheme name that refers to a specification for assigning identifiers within that scheme e.g. http:, ftp:, mailto:, file:…
The URI scheme defines the namespace
Documented in informational RFC (e.g.: RFC 3986)

“Decision Tree”

http://www.ncdd.nl/en/pid