Empowering the Usage of Persistent Identifiers (PID) in Local Research Processes by Providing a Service and Integration Infrastructure

Florian Krämer, Marius Politze, Dr. Dominik Schmitz
Content

• A quick word about RWTH Aachen University and Research Data Management
• Challenge: How to bring PIDs to researchers?
• Our solution: flexible web services
  • PID concept
  • Pre-existing solutions: support for the core processes teaching, learning and research
  • The web services for PID usage
  • Managing rights to update PID information
• Summary and Outlook
Empowering the Usage of Persistent Identifiers (PIDs)  |  Florian Krämer  |  Marius Politze  |  Dr. Dominik Schmitz  |  02.09.2016  |  RDA Europe  |  Views on PID Systems  |  Workshop

RWTH Aachen University and Research Data Management

- Project Group
  - IT Center
  - University Library and
  - Department Research and Career
  as central service providers

- Central services
  - Backup/Archive, Storage
  - Publication Server
  - Support
  - Training
Empowering the Usage of Persistent Identifiers (PIDs) | Florian Krämer | Marius Politze | Dr. Dominik Schmitz | 02.09.2016 | RDA Europe | Views on PID Systems | Workshop
Challenge: How to bring PIDs to researchers?

IT environments & researchers’ workflows

• The IT environments differ significantly
  • Historically grown local infrastructures
  • Diverse methods and machinery require specific IT solutions
  • More than one solution for the same problem

• PIDs need to be easy to use
  • Providing an easy to use web interface
  • Integration in their existing environment

• Researchers need to be aware of the existence of a PID system
Our solution: flexible web services

PID concept

- Early adoption:
  - PIDs should be registered as soon as data is produced
  - Even if it is not clear yet whether the data has any value
- Flexible usage to make data findable internally or also externally
- Minimal information stored with PID
  - URL
  - DATAURL
  - METAURL
  - RELATED ID
- PIDs to link information from and to different systems providing more (context) information on the data
- Using the Handle system, specifically ePIC
  - Simple registration process compared to DOI
  - Single PID prefix for the whole university

- Who is authorized to register or update a PID?
Our solution: flexible web services

Pre-existing solutions: support for the core processes teaching, learning and research

- Connect legacy systems with a consistent API
- An SOA that resembles university processes
  - Started with E-Learning
  - Generalize and try to apply to other fields:
    - Campus Management, Identity Management
    - Research Data Management / E-Science
- Security by design
  - Confidentiality
  - Integrity
  - Availability
- Protect personal and confidential data
Our solution: flexible web services

The web services for PID usage

• Create PIDs
  • Using OAuth2 for authorization
  • Assigned to the person and institute

• Create Formalized Metadata (RDF)
  • Based on Metadata schemas
  • Can be stored locally or in a centralized DB

• Display PID information and Metadata
  • Landing page for data
  • Contact information to acquire access

• Limit possible operations
  • Only a single PID generator
  • No delete operation
Our solution: flexible web services

The web services for PID usage
Our solution: flexible web services

Managing rights to update PID information

- Responsibilities are often shared
  - Multiple employees (e.g. admin and researcher)
  - Several organizations (e.g. Institute and Library)

- Lightweight way of passing permissions
  - One Time Access Tokens (OTA)
  - Passed between systems using the REST API

- Direct access to PID server for “trusted systems”
  - OTAs are self-contained
  - signed JSON Web Tokens with a shared secret

- Currently used to connect
  - Metadata Management
  - Archive
  - Publication Server
Summary and Outlook

- Diverse IT environments
- Simple PID concept
- Goal: integration of PID service in local IT solutions
- Usage of web service
- Decentralized rights management

- Connect additional (decentralized) systems
- Enable metadata harvesting by a central registry
- Enable metadata based search and retrieval while respecting the access rights