Webinar
Persistent IDs: Best Practices

Mark Leggott, Executive Director | November 29, 2016

Let's connect: mark.leggott@rdc-drc.ca | @mleggott

RDC DRC

Research Data Canada – Données de Recherche Canada

rdc-drc.ca | @rdc_drc

Research Data Canada is supported by CANARIE, an organization dedicated to advancing Canada's knowledge and innovation infrastructure.

> Unique Identifiers:
Current Landscape and Future Trends

Mark Leggott, Kathleen Shearer, Chantel Ridsdale

Download Link

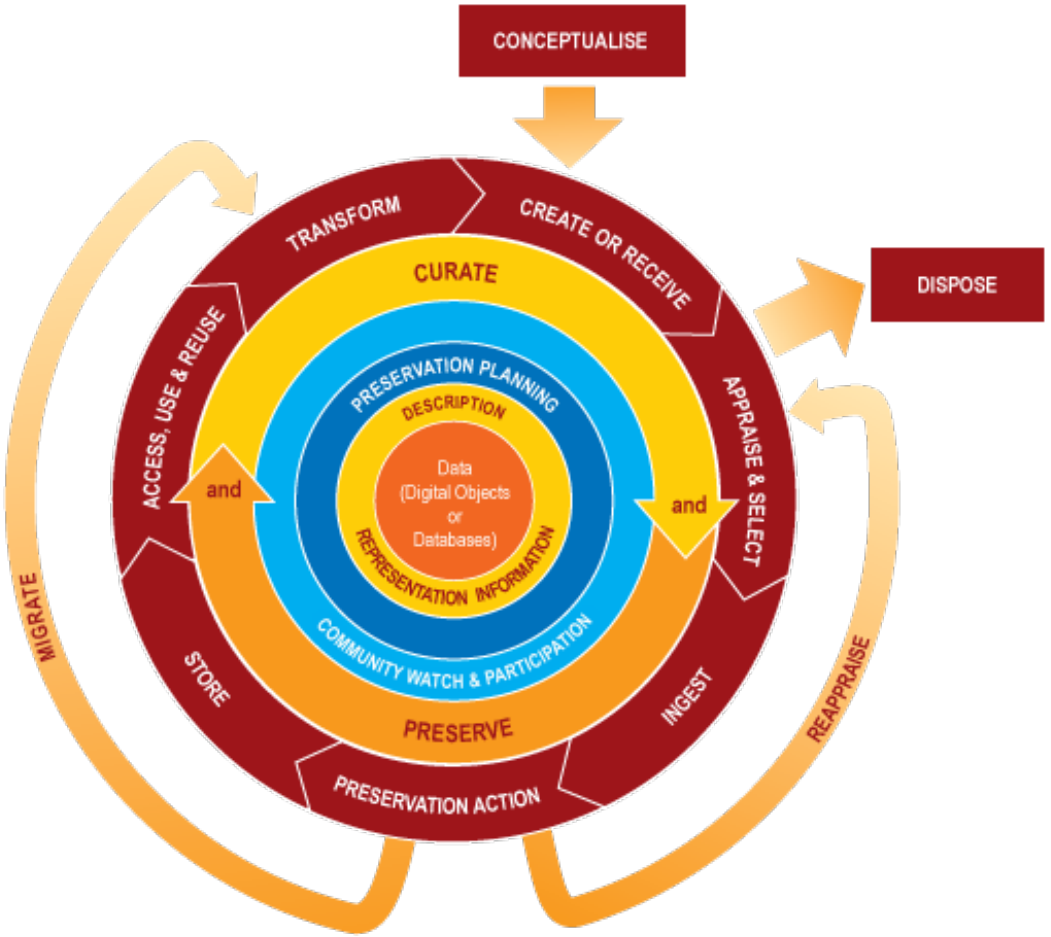
Research Data Canada works with stakeholders to ensure research data is available to support innovation that benefits all Canadians.

Universities

Federal Funding Agencies

Federal Research Agencies

Provincial Funding Agencies



Provincial Research Agencies

Open Data Organizations

Non-Profit & NGO Research Organizations

Commercial Research Organizations

International Agencies and Collaborators

The DCC Curation Lifecycle Model: http://www.ijdc.net/index.php/ijdc/article/viewFile/69/48.

> Vision

Canadians have access to national and international collections of publicly-funded research data.

> Mission

Work with stakeholders to ensure research data is re-usable in support of innovation that benefits all Canadians.

> Strategy

Development of a collection of integrated National Data Services that support researchers through all stages of the research lifecycle and promote compliance.

Best Practices Series

- New Document Series intended to provide guidance & recommendations on best practices for research data management
- Documents produced by Working Groups, which get direction from Committees and ED
- Anticipate 3-4 per year

Definitions

- Identifier (ID): provides a unique label for an Object
- Persistent ID (PID): provides a long-lasting reference to an Object, regardless of that Object's location
- PID System: provides a framework (i.e. software) facilitating discovery and access to Objects
- PID Minting Service: provides a framework (i.e. software) for creating unique PIDs

Identifiers

- Key goal is to uniquely describe each type of input/output in the research lifecycle to facilitate:
  - Discoverability
  - Citability
  - Linkability
  - Preservation
  - Quality control
- Standard and ubiquitous ID approach is key to creating a rich research data/information discovery system at the national level in Canada

Researchers

- RIDs: Researcher IDs
- Many systems (e.g. a Research Information System, the Canadian Common CV) use their own internally generated RID
- Facilitates unambiguous identification of a researchers and by association, their outputs
  - When researcher changes name
  - When researcher changes institution
  - Regardless of type of output

ORCID

- Open Researcher and Contributor ID
  - ORCID Organization maintains ORCID system
  - ORCID Identifier is the unique PID
  - ORCID Registry provides access to ORCID IDs
- Researchers controls what goes into their record
  - Information, Education, Employment, Funding Works
  - External access via ORCID site, API, 3rd party system integration
- Potential “glue” that binds researchers and their outputs regardless of where the digital asset is

ORCID CA

- Effort to create an ORCID Canada Consortium
  - ORCID CA Implementation Group moving ahead to define details
  - CRKN providing business lead; informal request made to CANARIE to provide technical support
  - Goal to launch for January 2017
  - Website has more information
- Will provide best level of access for organizations at the best price
- Goal is to have as many Canadian researchers as possible create ORCID IDs

Objects

- OIDs: Object Identifiers
- Many systems assign a unique ID in metadata to described objects – too many to count
- Various standard PIDs for Objects: PURL, Handle, ARK, GUID, DOI, etc.
- Which OID system used can often be an institutional decision, or determined by digital asset management software being used

DOI

- DOI: Digital Object Identifier
  - Based on Handle system
  - Useful and de facto standard for describing traditional outputs like journal articles
  - Also describes elements like images/SW – not common
- DataCite Canada is national DOI provider
  - Register to obtain/manage DOIs
- UBC DOI Minting software
  - Works with DataCite Canada to create one-off or mass DOIs along with metadata records

ARK

- ARK: Archival Resource Key
  - Developed by California Digital Library
  - Can be used to provide IDs for any type of object, digital, physical, living or intangible
  - Useful for describing individual digital elements, e.g. within a journal article (e.g. tables, images)
  - System provides metadata record along with the ID
- ARKs for research data recommended by ANDS

Organizations

- OrgIDs: Organizational Identifiers
- Less clear as to leading system for OrgIDs
  - Ringgold – commercial OrgID system
  - GRID – Open system from Digital Science
  - ISNI: International Standard Naming Identifier
    - Currently describes over 575,000 unique organizations
- Decision will need to be made soon to facilitate a standard approach to OrgIDs
- A combination of ISNI with a de facto Registry will likely be solution – ISNI+

Equipment

- EIDs: Equipment IDs
- Identifies a piece of research equipment or sub-component
- equipment.data Project
  - Currently defining common record – UNIQUIP
  - Focus on equipment, current methodology
- MERIL Project
  - Currently used in EU to describe research projects, but looking at how to describe key research equipment

Other

- Listing of additional ID approaches for
  - Specimens (geo, bio)
  - Chemicals, molecules
  - Other Objects
- More domain-specific in efforts
- Key to long-term goal of providing unique identifiers to any input/output in research

Key Recommendations

- Commit to the adoption of best practice PIDs
- Join the ORCID-CA consortium to develop a comprehensive national solution
- Require researchers to have an ORCID all levels
- Integrate ORCID into CCV
- Require DOIs for deposit in response to funder OA mandates (funder, institution, publisher, etc.)
- Adopt ISNI as standard for OrgIDs

RDC Actions

- Created PID Working Group to follow progress
- Develop communications strategy to facilitate broad adoption of standard PIDs in Canada
- Engage with the international research data management community on PID efforts
- Facilitate development of ORCID CA
- Promote use of PIDs by Canadian publishers
- Promote use of DataCite DOIs and UBC tools

RDC DRC

Research Data Canada – Données de Recherche Canada

rdc-drc.ca | @rdc_drc

Contact me:

mark.leggott@rdc-drc.ca | @mleggott

Research Data Canada is supported by CANARIE, an organization dedicated to advancing Canada's knowledge and innovation infrastructure.