Open Scholarship: Requirements & Resources

Open access, open data, open identifiers

Jay Brodeur
McMaster University Library
23-Feb, 2018

McMaster University LIBRARY

Itinerary

Open Access

- ✧ Tri-Agency policy
- ✧ Resources (MacSphere; DOAJ; Sherpa/ROMEO)

Research Data Management & Open Data

- ✧ Tri-Agency policy development
- ✧ Resources (DMP Assistant, OSF, Dataverse)

ORCiD ID

- ✧ What is it? Why use it?
- ✧ How to maximize its value.

McMaster University LIBRARY

Open Access

"Open Access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions"

- Peter Suber

Tri-Agency Open Access Policy on Publications

Government of Canada | Gouvernement du Canada

Home ➔ Collaboration between Federal Research Funding Organizations ➔ Policies and Guidelines ➔ Open Access

Tri-Agency Open Access Policy on Publications

1. Preamble

The Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Social Sciences and Humanities Research Council of Canada (SSHRC) ("the Agencies") are federal granting agencies that promote and support research, research training and innovation within Canada. As publicly funded organizations, the Agencies have a

http://www.science.gc.ca/eic/site/063.nsf/eng/h_F6765465.html

"Grant recipients are required to ensure that any peer-reviewed journal publications arising from Agency-supported research are freely accessible within 12 months of publication"

Open Access: Two venues for publishing

McMaster University LIBRARY

OA Journals ("Gold OA")

- Immediate OA provided by a publisher (processing fees may apply)
- Directory of Open Access Journals (doaj.org)
- Fees are eligible Use of Grant Funds

OA Repositories ("Green OA")

- Preprints and/or Postprints (before or after copy-editing)
- Directory of Open Access Repositories (opendoar.org)
- SHERPA/RoMEO: Database of publisher copyright policies

openaccess.mcmaster.ca

OA information for researchers

Customized for McMaster

Links to the Library's
Open Access Interactive Tool

https://www.youtube.com/watch?v=BPSq_zM5W80

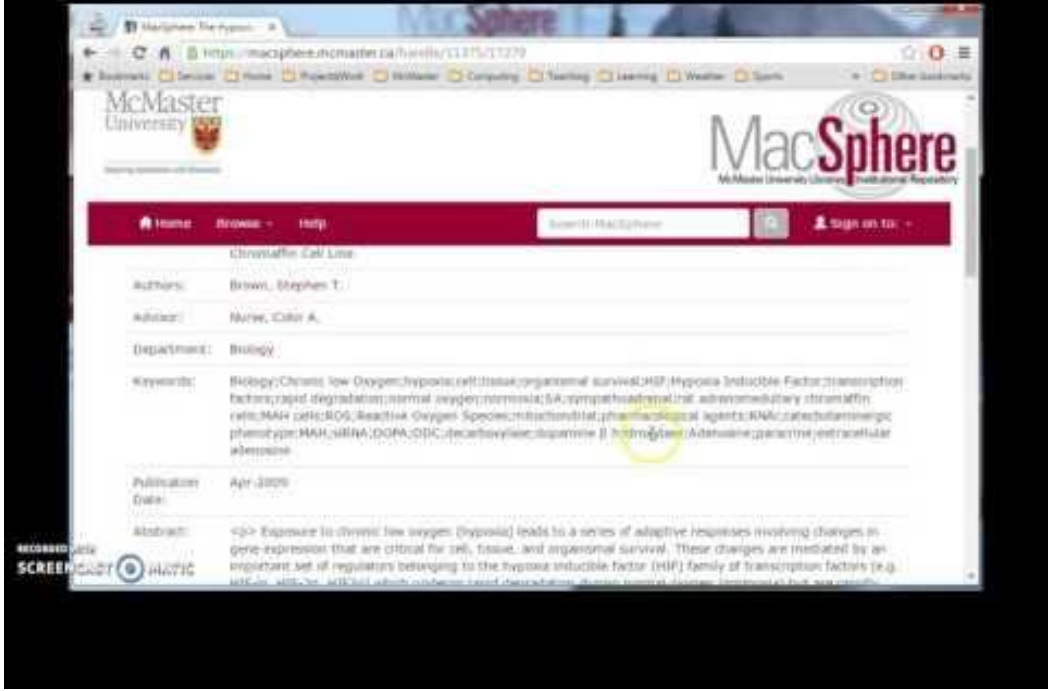
macsphere.mcmaster.ca

McMaster's institutional repository (IR)

Preservation and sharing for scholarly output at McMaster

Accepts many formats and outputs (theses, journal articles, data, figures, slideshows, etc.)

https://www.youtube.com/watch?v=Zrt5zHZfRk4



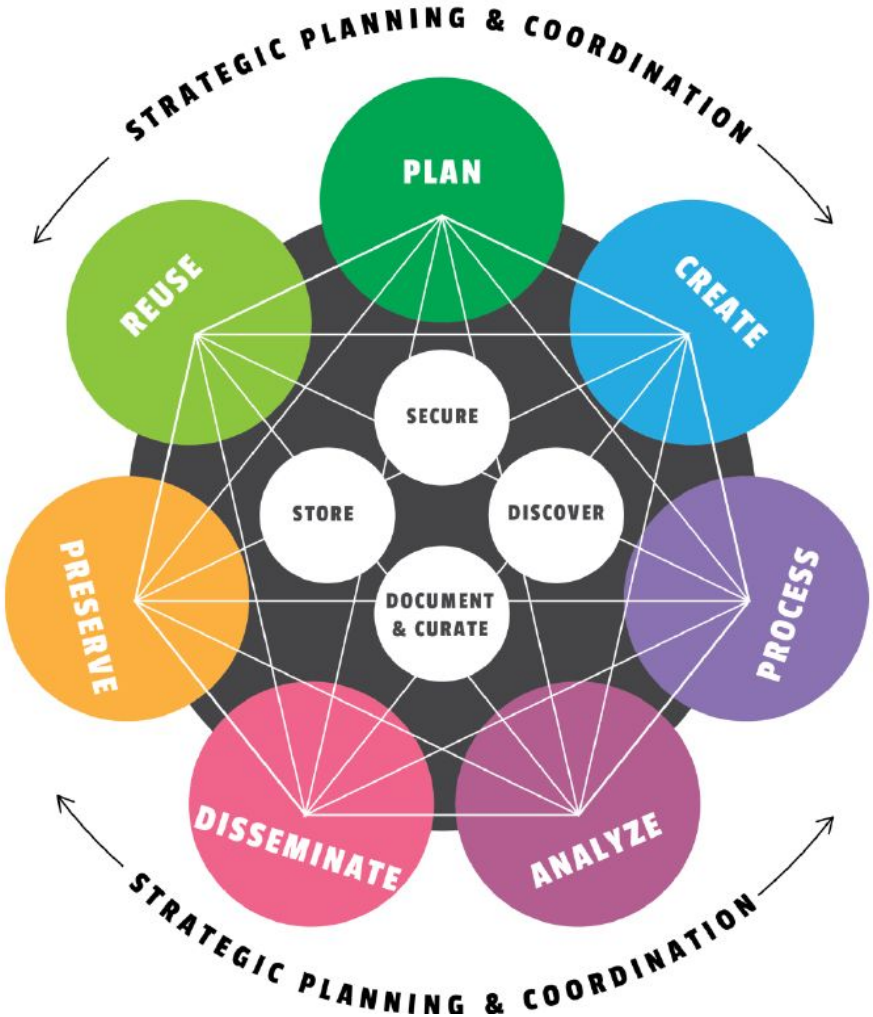
McMaster University LIBRARY

Open Data / Research Data Management

“Open data and content can be freely used, modified, and shared by anyone for any purpose”

- Open Knowledge International

Research Data Management is the active organization & maintenance of data throughout the research cycle to ensure data security, accessibility, usability, and integrity



The research data lifecycle

Courtesy: Leadership Council for Digital Research Infrastructure (LCDRI) Data Management Position Paper, 2017-08-31

Applying RDM best practices will benefit...

Researchers and their collaborators

- ✧ Improves research efficiency and productivity
- ✧ Provides extra credit for research work
- ✧ Increases research impact
- ✧ May help to meet funding requirements

Research Communities

- ✧ Accelerates discovery
- ✧ Enables validation and verification

Funders, governments and the public

- ✧ Improves return on investment
- ✧ Increases research transparency
- ✧ Data as a public good

Tri-Agency Statement of Principles on Digital Data Management
http://www.science.gc.ca/default.asp?lang=En&n=83F7624E-1

Tri-Agency Statement of Principles on Digital Data Management: **Expectations**

- **Data management planning**
- Constraints and obligations
- Adherence to standards
- Collection and storage
- Metadata
- **Preservation, retention and sharing**
- Timeliness
- Acknowledgement and citation
- Efficient and cost-effective

Tri-Agency Statement of Principles on Digital Data Management
http://www.science.gc.ca/default.asp?lang=En&n=83F7624E-1

Tri-Agency Statement of Principles on Digital Data Management: **Responsibilities**

Researchers

- incorporating best practices
- developing DMPs
- adhering to policies and standards

Research Institutions

- supporting best practices
- providing access to resources
- creating guidance and policies

Research Communities

- developing & promoting standards
- fostering excellence
- selecting repositories

Research Funders

- developing policies & guidance
- promoting data management
- providing peer reviewers

Draft Tri-Agency RDM Policy

- ✧ Tri-Agency draft data management policy — expected April, 2018
- ✧ 6-month consultation period; feedback will inform policy

Proposed policy — 3 possible requirements:

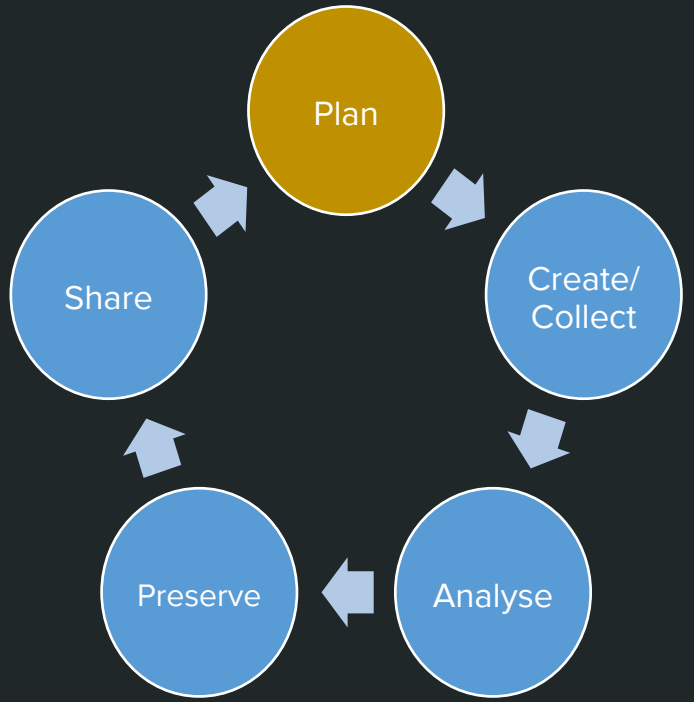
1. **Institutions**: Institutional Strategy
2. **Researchers**: Data Management Plans
3. **Researchers**: Data Deposit *(where appropriate)*

- ✧ Phased and incremental implementation

Link to M. Lucas' slides — Portage Network Event (30-Jan, 2018)

Data Management Resources

I. Planning for RDM



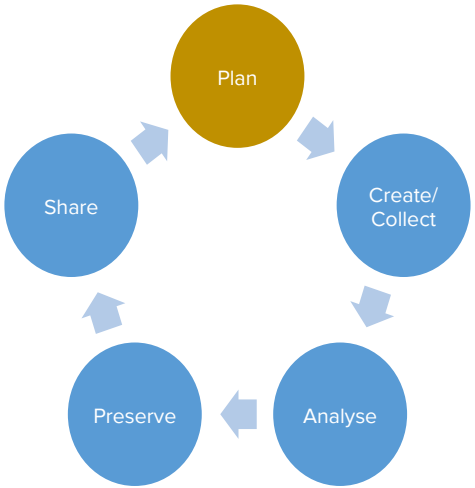
...it all starts with a (good) plan

A research Data Management Plan (**DMP**) should:

- ✧ Describe how you will manage data through all stages of your research
- ✧ Communicate a strategy for creating share-worthy data products

A **good** DMP will**:**

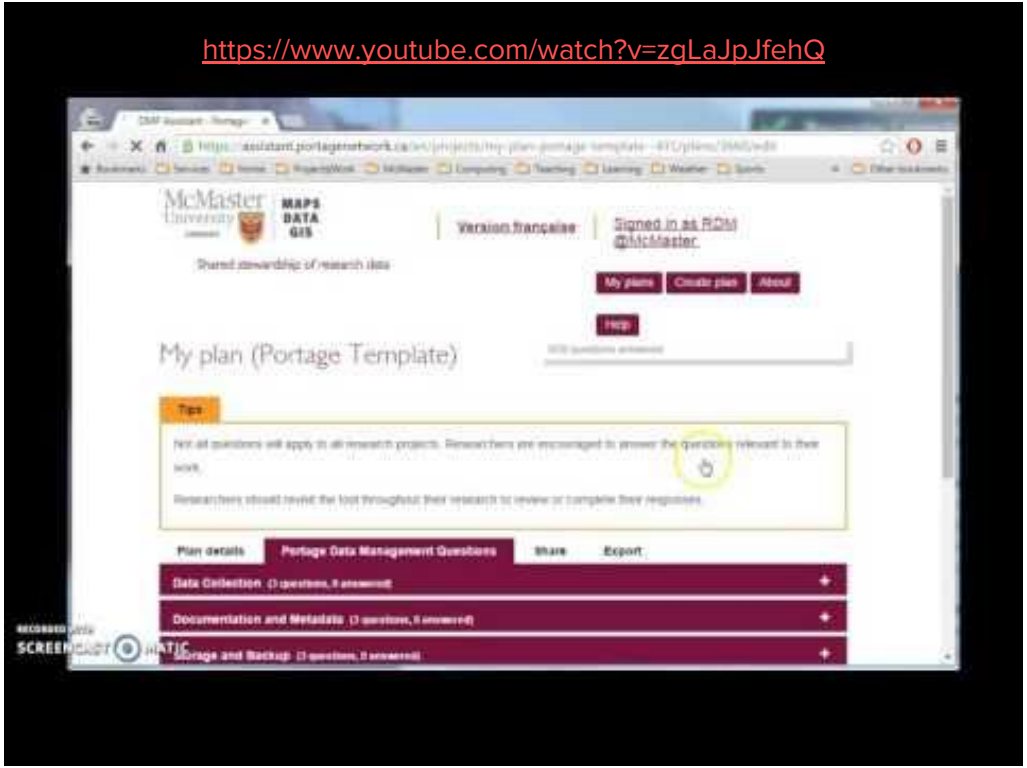
- ✧ Be completed at the time of study design
- ✧ Ensure compliance with policies / obligations
- ✧ Document and organize research activities
- ✧ Help identify support requirements
- ✧ *(Likely)* evolve with your study...



Portage DMP Assistant

- A web-based, bilingual data management planning tool.
- Available to all researchers in Canada.
- A guide for best practices in data stewardship.
- Exportable data management plans.

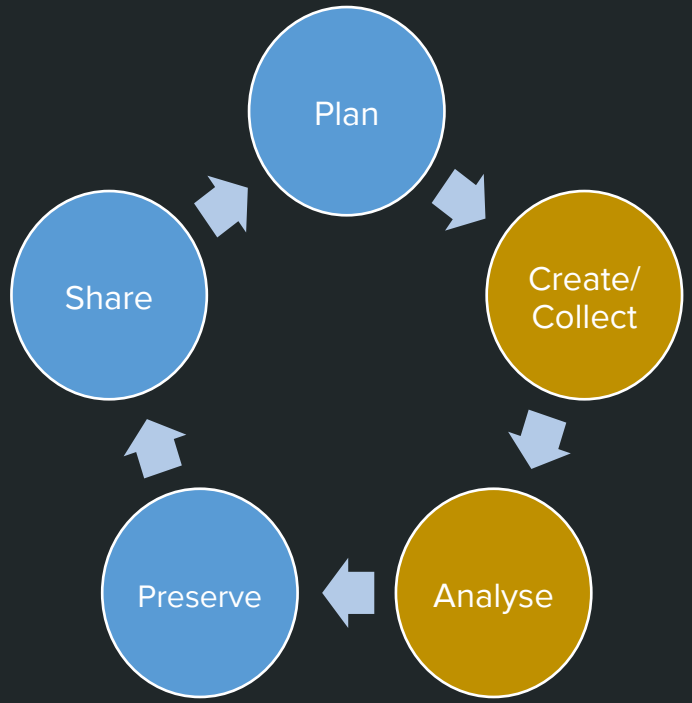
https://www.youtube.com/watch?v=zgLaJpJfehQ



https://assistant.portagenetwork.ca/

Data Management Resources

II. In-project data management



Things to consider...

How will you manage digital data in your research?

- What types of data will you collect and how much?
- How will you ***organize, secure***, *and* ***backup*** your data?
  - Are there ethical or commercial restrictions?

How will you provide access to collaborators?

- How will you ***describe*** your data so that others understand it?
- How will you control ***access*** to this data?
- How will you manage data versions?

3

2

1

McMaster University LIBRARY

3

copies of your data

2

1

3 copies of your data

2 copies are on-hand (easily accessible)

- 1 “**production**” (working) copy
- 1 “**production backup**” copy

1

McMaster University LIBRARY

3 copies of your data

2 copies are on-hand (easily accessible)

- 1 “**production**” (working) copy
- 1 “**production backup**” copy

1 copy is in another location (“off-site”), with a ***trusted*** service provide

McMaster University LIBRARY

3 2 1

“**Production**” copy ➜ Where you work with the data

- PC, laptop, mobile device, etc.

“**Production backup**” copy ➜ Easily accessible (+ versioning?) backup

- External hard drive with backup software
- MacDrive (seafile): https://macdrive.mcmaster.ca
- Dropbox, Google Drive, etc.

Off-site “Archived” Backup

Providers / Services:

- Campus / Consortium-hosted (RHPCS)
- Remote, Commercial
    - Backblaze, Iron Mountain, JustCloud, etc.

Open Science Framework

https://osf.io/

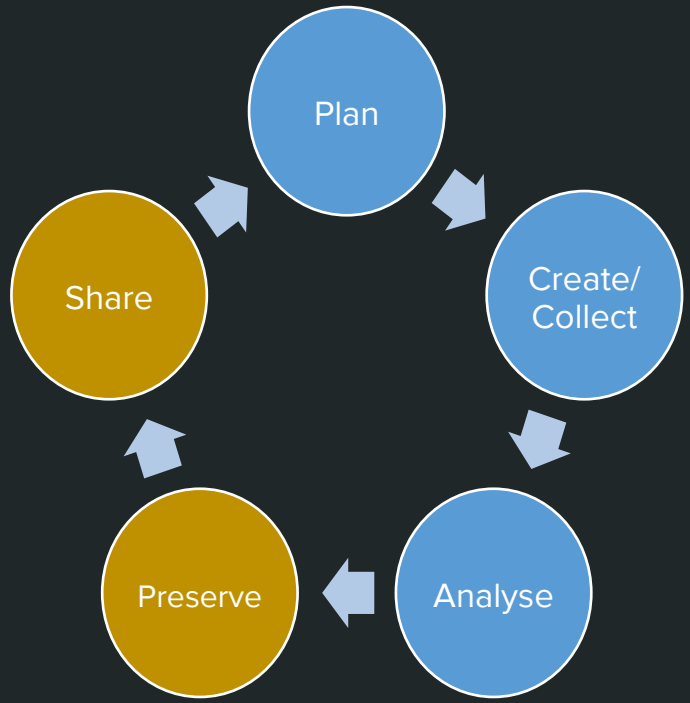
- Free, open source web application
- Connects and supports the research workflow
- Integrates with existing services (e.g. Google Drive, Dropbox, etc.)
- Supports collaborative workflows
- Facilitates data publishing to selected data repositories

https://youtu.be/2TV21gOzfhw

Workshop demo: https://osf.io/muq5f

Data Management Resources

III. Preserving and sharing data products



Things to consider…

How will your data products be stored in the long-term?

- How to ensure that it remains ***integral*** and ***secure***?
- Who will assume long-term ***responsibility*** for your data?

How will others access your data products?

- What data (if any) can/should be shared? Who should have access?
- How will you manage legal, commercial & ethical constraints?

How to maximize credit for sharing your data?

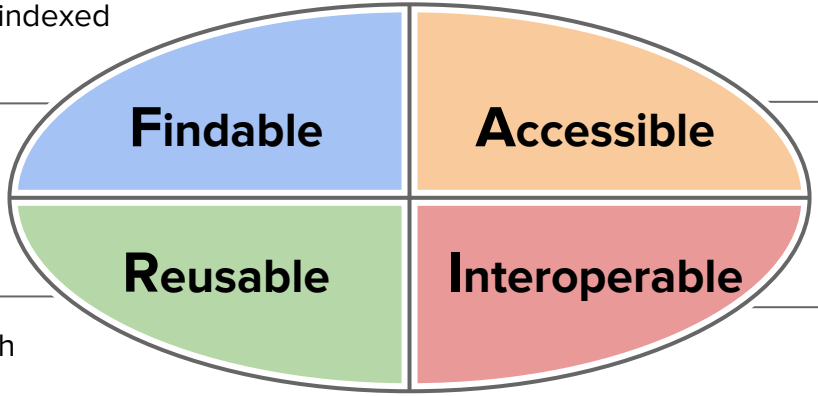
- In which repository should you deposit your data?
- How to ensure that your data is **FAIR**
  (*findable*, *accessible, interoperable* and *reusable)*?

The FAIR Guiding Principles

McMaster University LIBRARY

F1: (meta)data have a globally unique and eternally persistent identifier
F2: data are described with rich metadata
F3: metadata clearly and explicitly includes the ID of the data it defines
F4: (meta)data are registered and indexed in a searchable resource

A1: (meta)data retrievable by their ID using a standardized protocol
A1.1: protocol is open, free and universally implementable
A1.2: protocol allows for AuthT/ AuthZ where needed
A2: metadata is always accessible



R1: meta(data) richly described with accurate and relevant attributes
R2: (meta)data released with a clear and accessible data usage license
R3: (meta)data associated with detailed provenance
R4: (meta)data meet domain-relevant community standards

I1: (meta)data use a formal,accessible, shared, broadly applicable language for knowledge rep.
I2: (meta)data use vocabularies that follow FAIR principles
I3: (meta)data include qualified references to other (meta)data

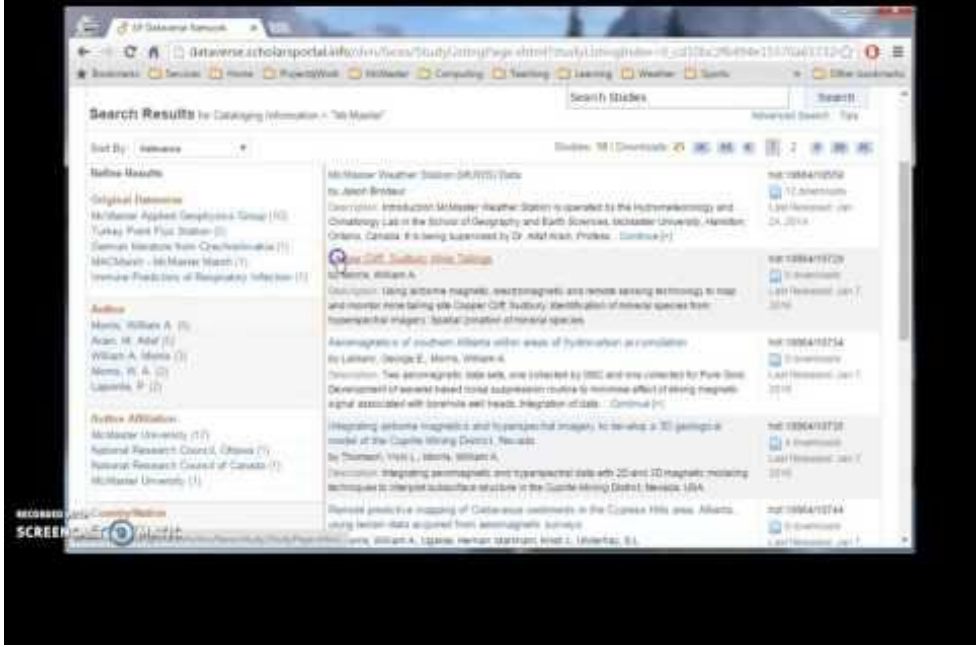
Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Bouwman, J. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific data, 3. ***doi:10.1038/sdata.2016.18***

Scholars Portal Dataverse

Scholars Portal
A Service of the Ontario Council of University Libraries

- A data repository for researchers at Ontario's universities.
- An online platform to share, preserve, cite, explore and analyze research data.
- Allows researchers to control how they share their data.
- Supports data DOI registration through Datacite Canada.

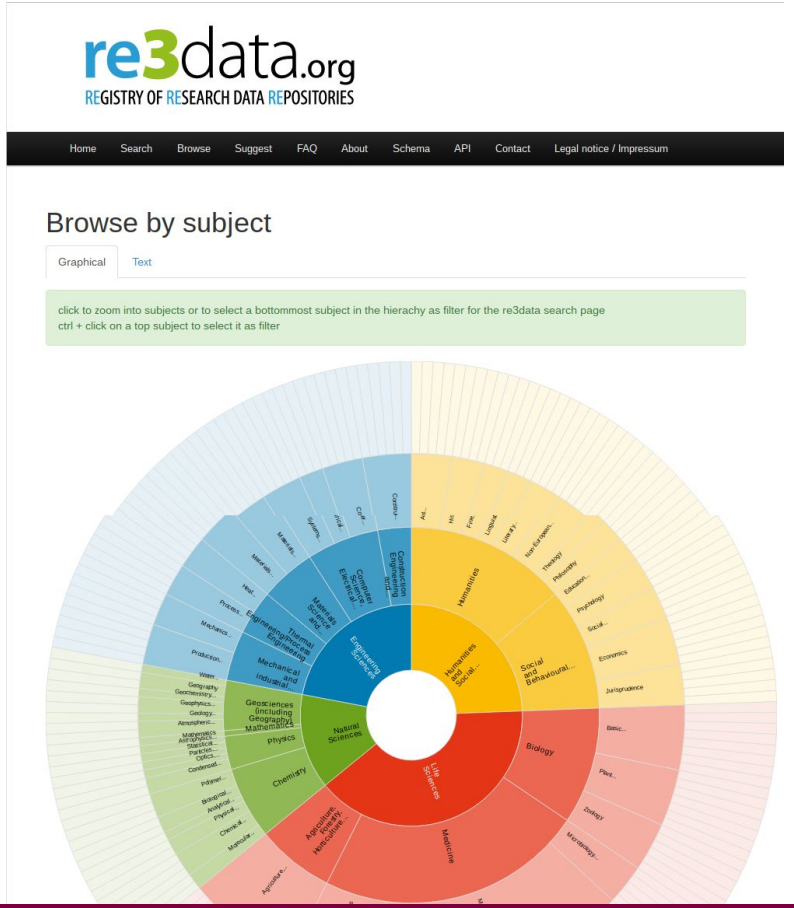
https://www.youtube.com/watch?v=UDFGqRY61fQ



http://dataverse.scholarsportal.info or explore the demo site: https://demodv.scholarsportal.info/

re3data.org

- Registry of research data repositories



McMaster University LIBRARY

ORCiD ID

What is an ORCID iD?

ORCID = Open Researcher and Contributor ID

- Unique identifier for use in scholarship
- Transparent system for linking researchers to:
  - activities & outputs;
  - individuals; and,
  - Organizations.
- National adoption underway in UK, AUS, NZ
- > 50 000 ORCID iDs registered from .ca domain

orcid.org

What does an ORCID iD do?

- Disambiguates researchers
- Enables connection outside of the institution
- Simplifies/automates data entry processes
- Improves data quality and reuse (machine-actionable data)
- Improves online presence of researchers' products

orcid.org

ORCID-CA

ORCID-CA is the ORCID Consortium in Canada, facilitated by ORCID and the Canadian Research Knowledge Network (CRKN)

Benefits:

- ✧ Premium ORCID membership at lower cost
- ✧ Access to Premium Member API = potential to integrate with University systems and workflows
- ✧ Dedicates Member Services Coordinator

orcid.org

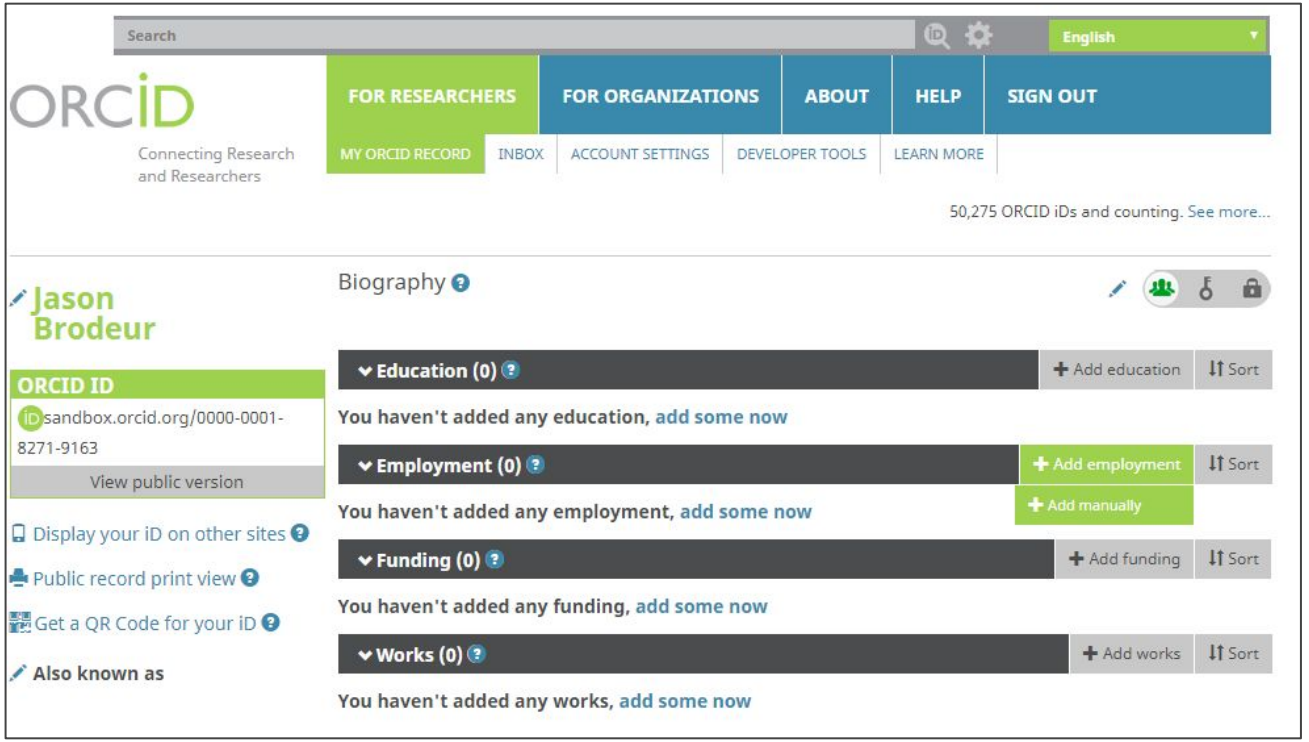
Affiliating an ORCID iD with McMaster

Current: Add McMaster University in ***Employment*** Section

Step 1

Sign up / Sign in at orcid.org

In Bio, select:
+Add employment
+Add manually



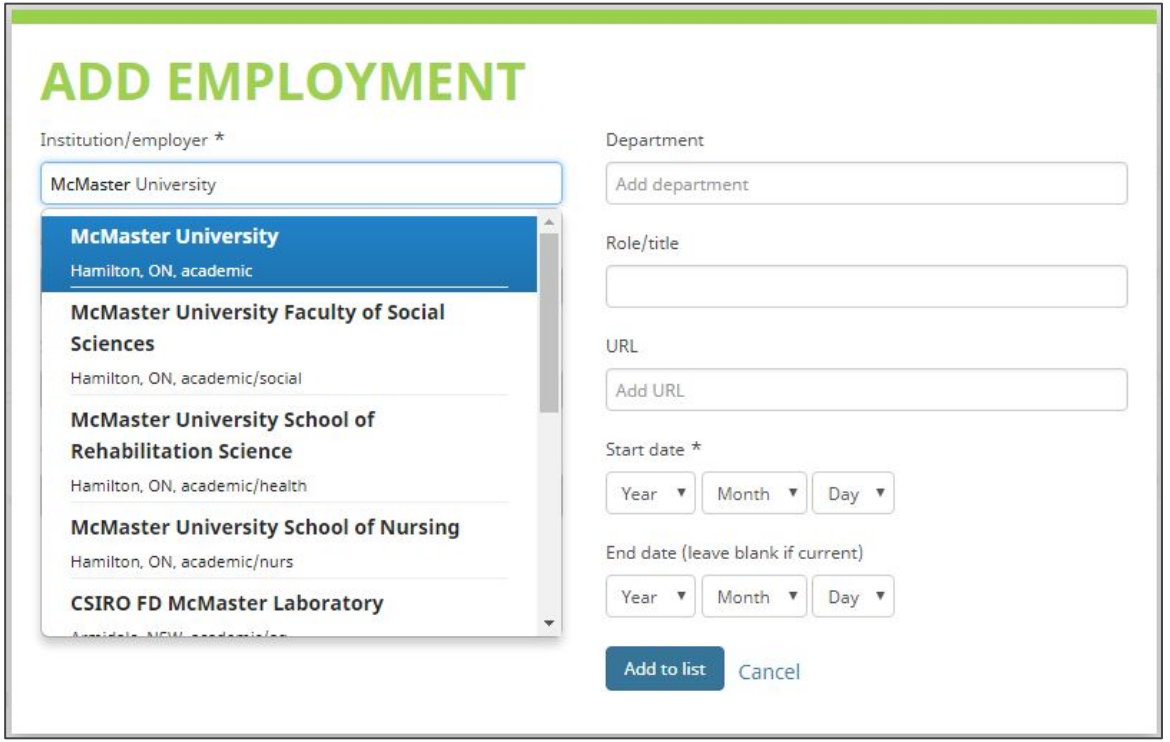
Affiliating an ORCID iD with McMaster

Current: Add McMaster University in ***Employment*** Section

Step 2

Enter ‘McMaster’ as Institution/employer

Select ‘McMaster University’ (or sub-unit) from drop-down list



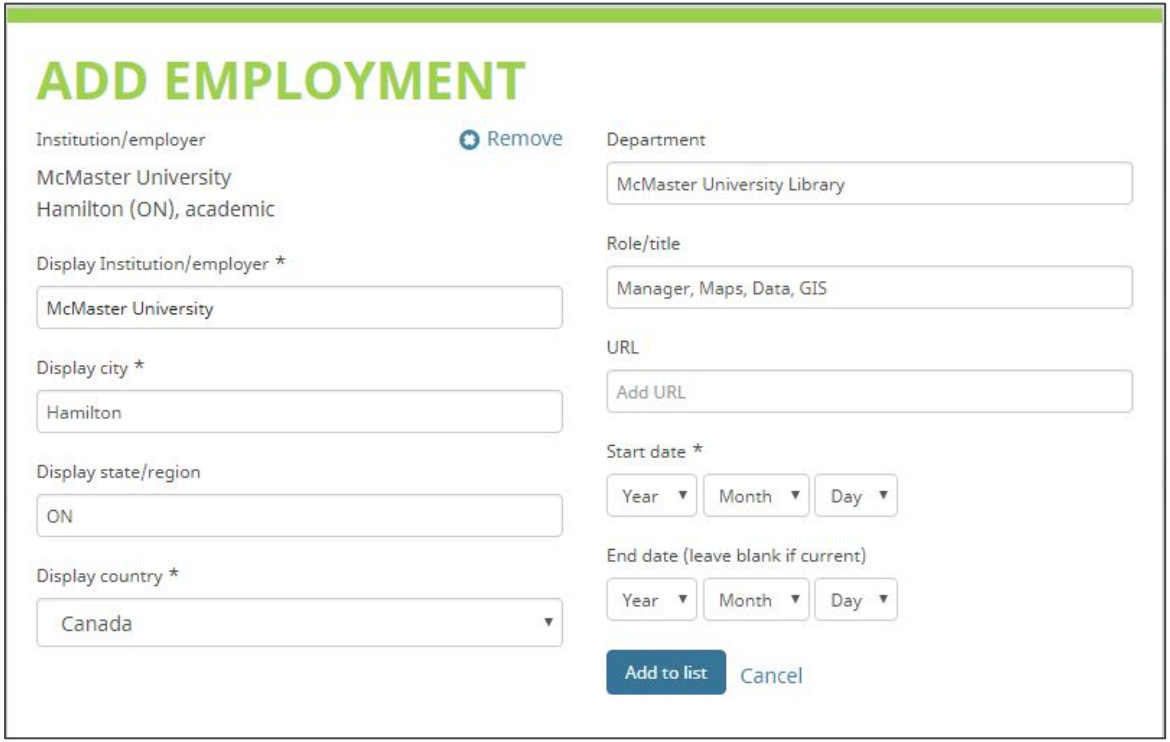
Affiliating an ORCID iD with McMaster

Current: Add McMaster University in ***Employment*** Section

Step 3

Add remaining position information (mostly optional)

Click “Add to list”

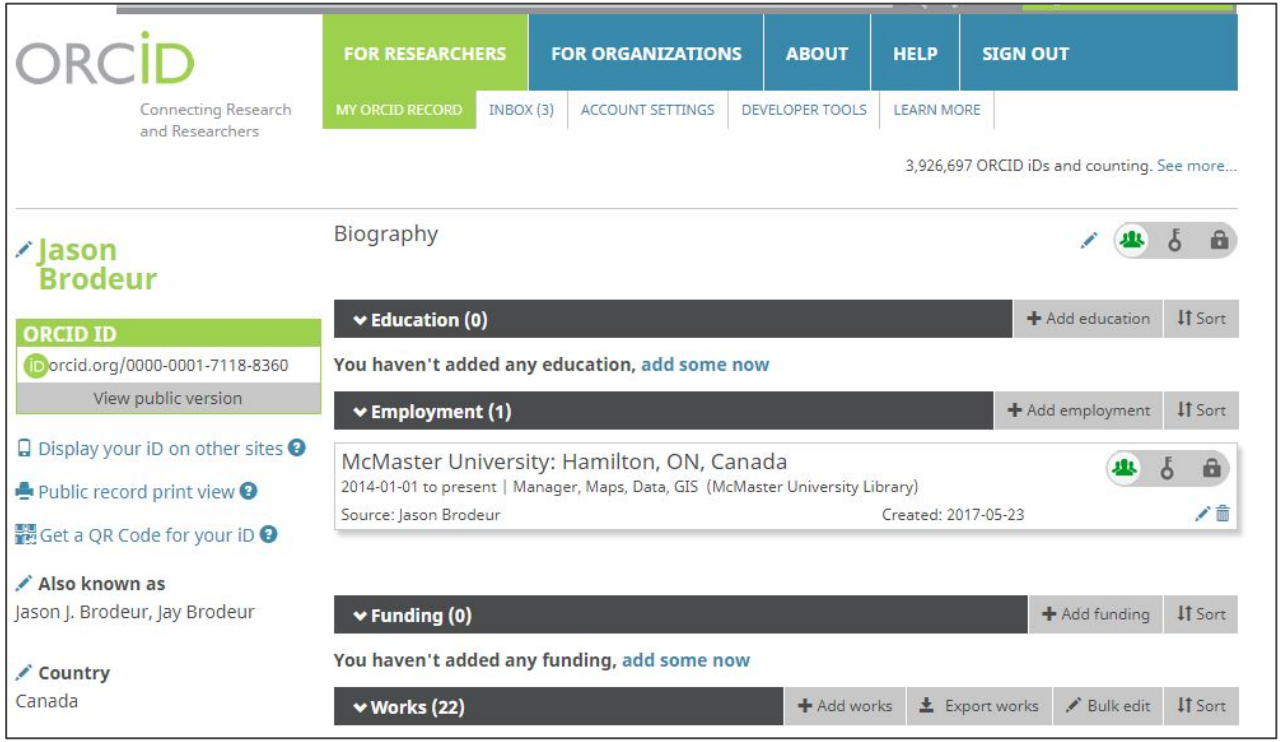


Affiliating an ORCID iD with McMaster

Current: Add McMaster University in ***Employment*** Section

Profile will now show employment

Affiliation established



Affiliating an ORCID iD with McMaster

Coming Soon:

- ✧ Affiliation can be initiated by the institution through ORCID-CA membership API
- ✧ Researchers will be prompted to accept/reject affiliation with McMaster

More Information on ORCID and ORCID-CA

McMaster University LIBRARY

ORCID:

Learn more; register at: **https://orcid.org/**

ORCID-CA:

Introductory Webinar: https://attendee.gotowebinar.com/recording/3169306365980109057

Webinar slides: https://goo.gl/Oi70kY

Joint Statement of Principles: https://goo.gl/2vb5p0

More information on researcher profiles:

http://hsl.mcmaster.libguides.com/impact/profiles

Where to find more information

Open Access & Open Educational Resources:

- openaccess.mcmaster.ca
- macsphere.mcmaster.ca
- Olga Perkovic (perkovi@mcmaster.ca):
  Research and Advanced Studies Librarian

Open Data & Research Data Management

- portagenetwork.ca
- library.mcmaster.ca/rdm
- RDM@Mcmaster (rdmgmt@mcmaster.ca)

McMaster University LIBRARY

Thank You.

Jay Brodeur, brodeujj@mcmaster.ca

Research data are contents that are used as primary sources to support research, scholarship, artistic activity or research-creation, and that are used as evidence in the research process and commonly accepted in the research community as necessary to validate research findings and results.

Canadian Tri-Agency

Research Data Management is the active organization & maintenance of data throughout the research cycle to ensure its security, accessibility, usability, and integrity

Research Data Management (**RDM**) is...

... the active organization & maintenance of data

... the application of best practices to ensure data security, accessibility, usability, and integrity

... a set of activities resulting in self-describing data sets that can be discovered and reused