

Research Data Management at the University of Alberta

Geoff Harder, Associate University Librarian

Strategic Workshop for Research Data Management

Sponsored by Research Data Canada and University of Alberta

Council Chambers, University of Alberta

November 17, 2015

Background

Foundation:

Research Data Centre (2001)

Data Library (1992)

Research data expert(s)

Extensive network of partnerships and contacts
e.g. DLI (1996), DDI Alliance (2007), Canadian
Polar Data Network (2010)

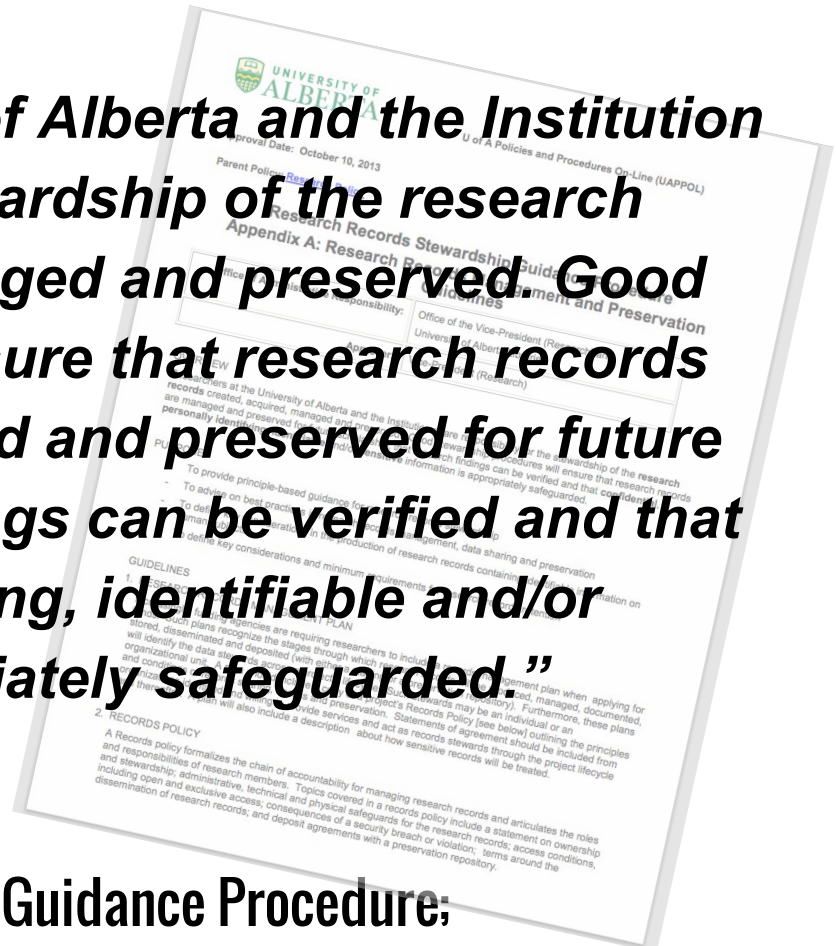


RDM: Moving from abstract problems to concrete solutions



- Where we are → where we want (or need) to be
- **Policies, infrastructure** and **people** will help us get there
- A **culture of data stewardship** will indicate we have arrived

“Researchers at the University of Alberta and the Institution share responsibility for the stewardship of the research records created, acquired, managed and preserved. Good stewardship procedures will ensure that research records [incl. research data] are managed and preserved for future scholarship, that research findings can be verified and that confidential, personally identifying, identifiable and/or sensitive information is appropriately safeguarded.”



Policy: Research Records Stewardship Guidance Procedure;
Appendix A: Research Records Management and Preservation Guidelines (2013)

Infrastructure

To support:

Organization

Access

Sharing
(Credit)

Archiving

RESEARCH DATA MANAGEMENT

Organize my data

How to prepare a data management plan & protect your data

Share my data

Why you would share, data sharing tools, metadata & how to get credit

Access existing data

How to locate data and access data tools, terms & citations

Archive my data

How to find a repository, prepare your data & submit your data

Good data management is informed by funder and institutional policy and encompasses the full data lifecycle including collection, analysis, dissemination and potential reuse.

research data management week 2015

[#uardmw15](#)

DMP ASSISTANT



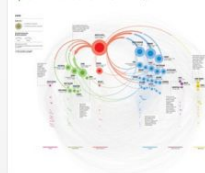
DATAVERSE



Tweets



Small Multiples 12 Nov
@smallmultiples
New project for Nature Index - Explore how the world's scientists collaborate natureindex.com/country-output... #NICollab15 pic.twitter.com/2bOkw1Q37z



People

Gifts and Donations

Copyright

Contact Us

Policies

Our Collection

Employment

Search this Site

Tools: DMP Assistant (formerly DMP Builder)



Grant Assist Program (Health Sciences)



Version française

Home About Help

DMP Assistant is a bilingual tool for preparing data management plans (DMPs). The tool follows best practices in data stewardship and walks researchers step-by-step through key questions about data management.

Step 1 Sign up with DMP Assistant

Step 2 Sign in and select a template under Organizations. The Portage template is the default.

Step 3 Answer the questions that are relevant to your work. Guidance and examples are provided.

Step 4 Revisit the tool throughout your research to review or revise your answers.

Sign in

If you have an existing account with DMP Assistant or previous version of DMP Builder.

Sign up

New to DMP Assistant? Sign up today.

Please note that we are currently working on single sign-in authentication. For now, please create a new DMP Assistant account. You will have the option to link your DMP Assistant account to your campus ID when that feature becomes available.

About Us

Changes to CIHR's Open Programs

Feedback Processes

Workshops

Archived Workshops

- Budget
- Canadian Common CV
- Finding Funding
- Foundation Competition
- Grant Application Process
- Letters of Support
- Non-OOG Funding Opportunities
- Peer Review Committee Selection
- Proposal Writing
- SPOR SUPPORT Unit

Data Management Plans

- Bibliometrics & Librarian Support

Home > Workshops > Archived Workshops > Data Management Plans

Data Management Plans

The Value of a Data Management Plan to Your Research: get a head start! This workshop discussed how a management plan can assist you in getting a head start in your next research project and learn about [Data Management Plan \(DMP\) Builder](#), an online Web-tool for drafting such plans. Increasingly funding agencies expecting researchers to have a data management plan when they submit their grant applications. This workshop reviewed the steps in preparing a plan and how you can get the most out of it. (Date Delivered: March 2014)

A web application for data management planning UNIVERSITY OF ALBERTA LIBRARIES

C. Humphrey, D. Borth, N. MacGregor, K. Roun, T. Jenkins, T. Chatterley, W. Shi, K. Roark, G. Gustin, G. Harder
University of Alberta Libraries, 401 Common Library, University of Alberta

Introduction
There is a growing call for researchers to prepare data management plans (DMPs) for their research projects. DMPs describe how you will collect, store, manage, share and disseminate your data throughout the course of a research project. They outline how data will be handled after a research project has completed.

Using the DMP Tool
After creating an account and logging in, users will be able to create a new plan. They choose a template from a list of available templates and then answer a series of questions to create a customized plan for their own data.

Conclusion:
The DMP tool will allow researchers to quickly and efficiently create DMPs that address specific funding and institutional requirements. These DMPs will be available to researchers and librarians and can be used as a template for other researchers.



Implemented DMPonline software for UA community in July 2014 with an open offer for others to make use of the tool

Toolkit - Dataverse

UAL DATAVERSE NETWORK

Welcome to the University of Alberta Libraries' Dataverse network. This service is provided to help researchers **publish, analyze, distribute** and **preserve** their data and datasets.

DATAVERSES

31 Dataverses

A **Dataverse** is a container for research data studies, customized and managed by its owner.

RECENTLY RELEASED DATAVERSES

Canadian Heart Health Survey	Oct 15, 2015
Energy Transitions in Canada	Oct 13, 2015
Mapping and Expanding Library Classification for Marginalized Domains: Feminism as a Prototype	Oct 13, 2015
Public Attitudes in Alberta Toward the Formerly Mentally Ill	Oct 13, 2015
University of Alberta Gradwand Survey	Oct 13, 2015

[View More >](#)

STUDIES

225 Studies, 2,475 Files, 1,409 Downloads

A **study** is a container for a research data set. It includes cataloging information, data files and complementary files.

RECENTLY RELEASED STUDIES

Alberta Survey, 2005 by Population Research Laboratory	Nov 12, 2015
Alberta Survey, 2007 by Population Research Laboratory	Nov 12, 2015
Edmonton Area Survey, 1985 by Population Research Laboratory	Nov 12, 2015
Edmonton Area Survey, 1986 by Population Research Laboratory	Nov 12, 2015
All Alberta Survey, 1987 by Population Research Laboratory	Nov 12, 2015

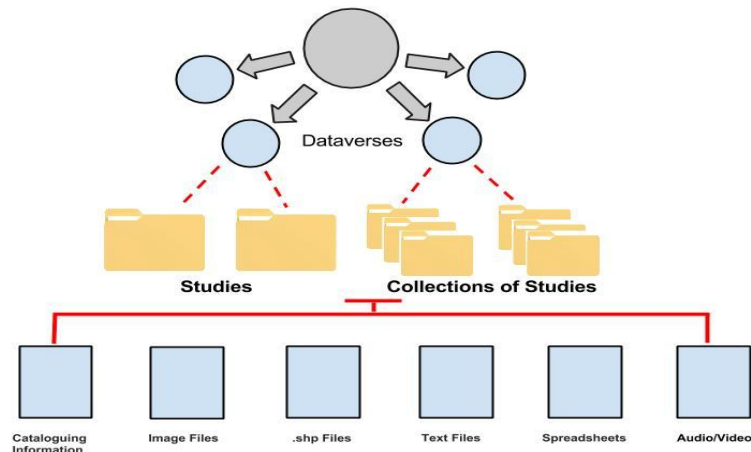
[View More >](#)

MOST DOWNLOADED STUDIES

Gamegate Reactions by Rockwell, Geoffrey; Suomela, Todd	619
Obispos y Prebendas de Chiapa	141
Addresses, Material Type and Heights for City of Edmonton Buildings in 1913. by Williams, Jaron ; Laliberte, L.W.	79
Open Genetics Lectures (OGL) Fall 2015 - Individual Chapters in .docx format by Locke, John; Deyholos, Michael; Harrington, Michael; Carham, Lindsay; Kang, Min	70
CGRI 1961 Census of Canada: Aggregate data and geography = CGRI 1961 Recensement du Canada: Données agrégées et cadre géographique by Gaffield, Chad; Baskerville, Peter; St-Hilaire, Marc; Amrhein, Carl; Cadigan, Sean; Moldofsky, Byron; Darroch, Gordon; Bellavance, Claude; Normand, France	70

[View More >](#)

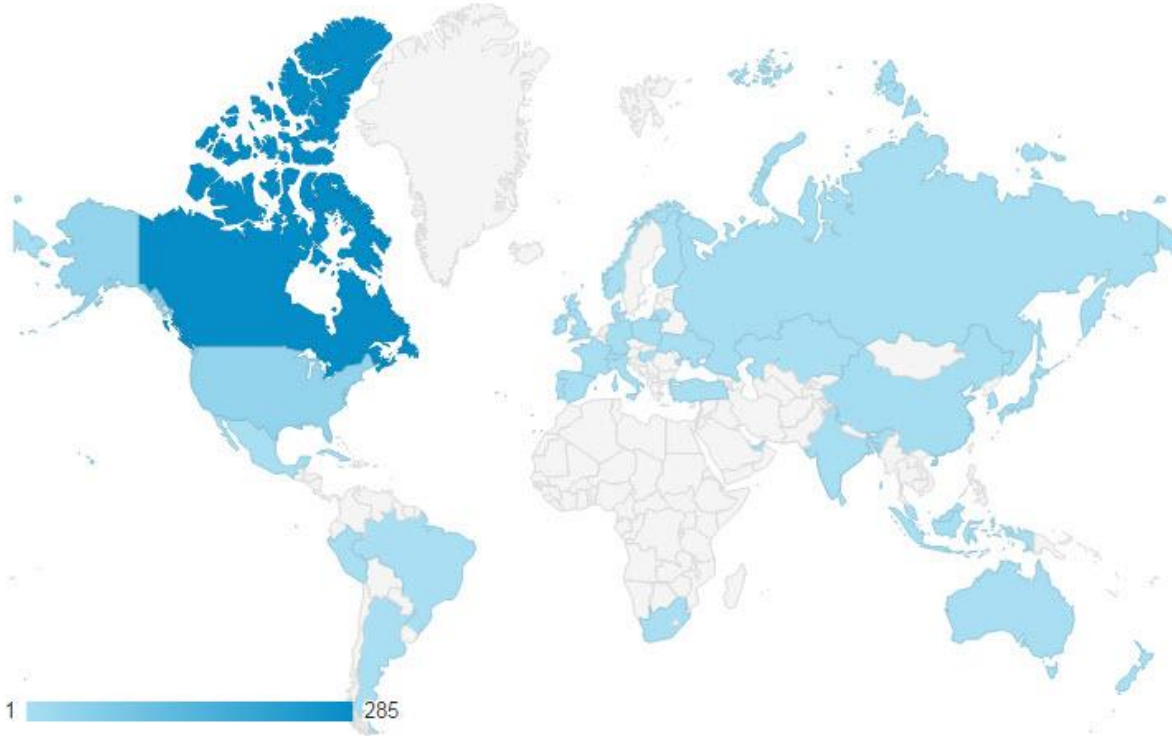
Dataverse Network



Adapted from : <http://www.dlib.org/dlib/january11/crossas/01crossas.htm>



Launched in March of 2014, the Univ of Alberta's Dataverse is one of three implementations in Canada (see UBC and Scholars Portal/OCUL). Dataverse represents an example of a growing global network.



HARVARD UNIVERSITY HARVARD.EDU

IQSS
The Institute for Quantitative Social Science

CONTACT US OPPORTUNITIES

Search

The IQSS Story

About IQSS • Teams & Programs • People • News & Events

Understanding and solving society's greatest challenges
through better, faster, bigger, and more collaborative social
science. [Learn more](#)

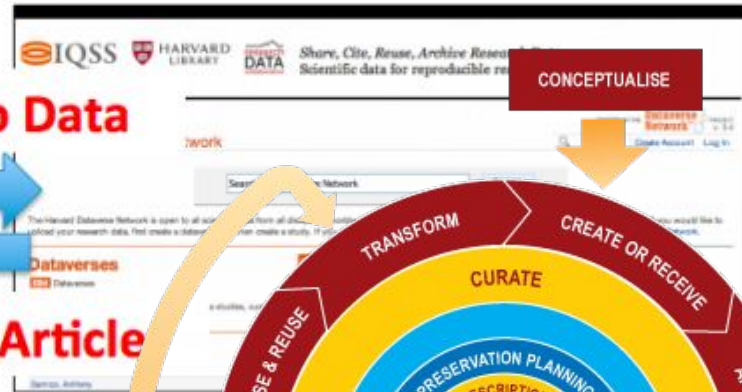
Helping social scientists leverage advances in data science, computing technology, quantitative methodology, and behavioral research, through activities in

Global view of locations where users are accessing the UAL Dataverse repository.

PKP's Open Journal System (OJS)



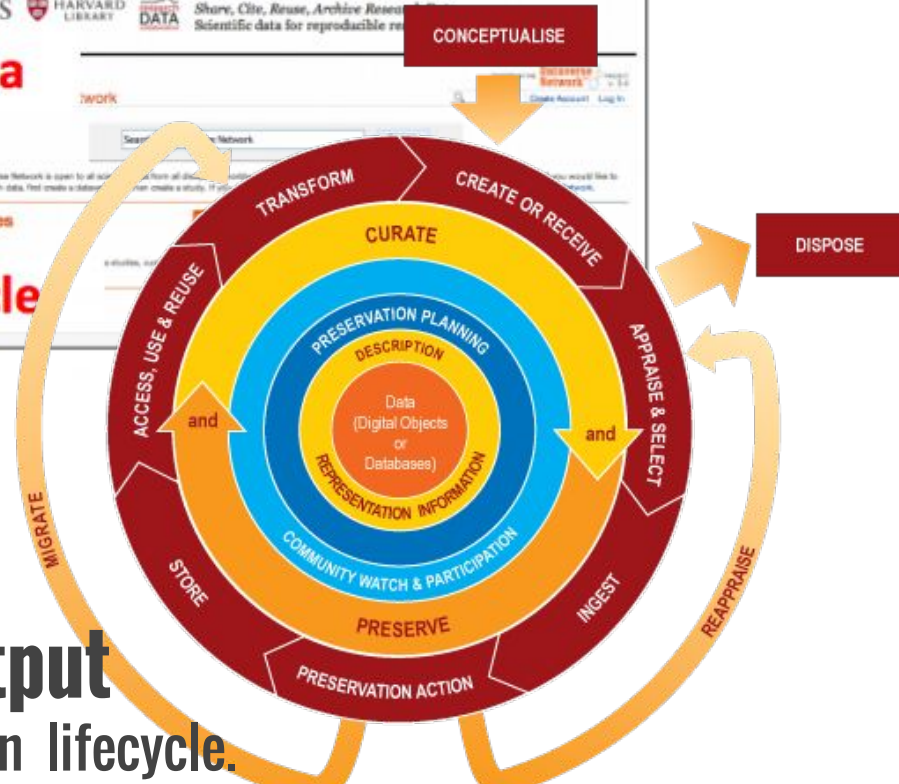
Harvard Dataverse Network



Citation to Data



Citation to Article



Data as a **first class scientific output** in the research and scholarly communication lifecycle.

Toolkit: Repositories

Keep it.

Share it.

Use it.



Get ahead
on your repository.

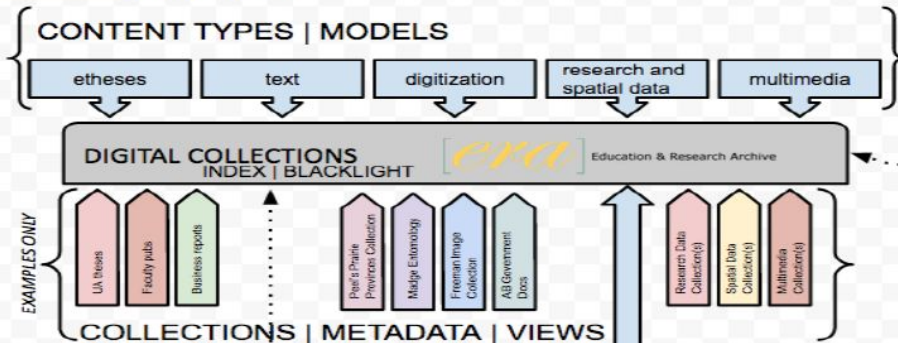
Multi-Purpose Repository Solutions
Flexible User Interfaces
Durable Digital Asset Management

Toolkit - Digital Asset Management System (Education and Research Archive repository)

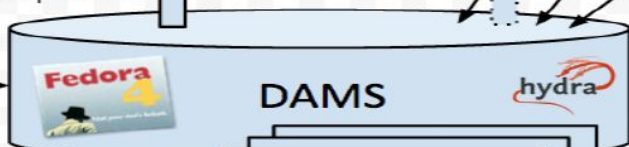


**HYDRA
NORTH DAMS
PROJECT
2014/15 1.0**

MAIN DISCOVERY INDEX (LIBRARY_ALL)



@archivematica



ETDs, books, papers, newspapers, reports, gov pubs, research data, images, maps, multimedia, etc...



USE CASE - EXHIBITS in Omeka

Characteristics:
project-oriented;
temporary lifespan: asset management lives in DAMS (accountability)

OER

Assets managed in DAMS but read/accessed elsewhere, e.g. Moodle

NOTES

- DAMS = Digital Asset Management System; connection to preservation storage, i.e. Swift + OpenStack
- Main discovery = Blacklight, not full text index; representation of some materials at collection level, e.g. research data, digitized newspaper title (not full text)
- Content models = e-thesis (multipart), book, journal, article, newspaper (article segmented), image (large, small), audio, video, data set, spatial data, etc.
- DAMS development driven by user stories
- Access to original and derivative files to support accessibility and digital scholarship, e.g. OCR, TIFF/JP2
- Archivematica may be used as a processing/packaging tool for certain types of content going into DAMS, e.g. research data, archival
- Utilize digital object identifiers (DOIs) or equivalent persistent IDs, where logical Respect metadata in its native format, whenever possible
- UJA Archives digital records to be included at later stage.

Health Research Data Repository (Nursing)



- **Secure virtual research environment (VRE)** for housing and managing both qualitative and quantitative health research data and metadata throughout their lifecycle
- Provides **data management guidance** and **educational opportunities** to researchers and their team
- Provides **secure transferring and remote access** to research data
- Provides a regular **suite of qualitative and quantitative analytic software**
- Supports and promotes **health research and multidisciplinary collaboration**
- **Clearly operationalized policies and procedures**, with user orientation process
- **Centralized transcribing service**
- Promotes **secondary use and re-purposing** of research data



Shifting Cultures - Research Data Management Week

RESEARCH DATA MANAGEMENT WEEK PLENARIES

The Perils and Prospects of Digital Scholarship in 21st Century Canada

Ted Hewitt: President, Social Sciences and Humanities Research Council

Monday May 4, 10am-11am - ECHA 1 498

Virtual Research Environments and the Management of Data: The Case of Semantic Collaborative Corpora Analysis

Christoph Schindler

Tuesday May 5, 12-1pm - ECHA 1 498

Large Data Transfers and Mirroring: Technologies and Their Implications

Paul Lu: Computing Science, University of Alberta

Wednesday May 6, 10am-11am - ECHA 1 498

Dealing with the Crowd: Managing Data from Citizen Science and Crowdsourcing Projects

Todd Suomela, Elisabeth Beaubien, Justine Kummer

Friday May 8, 1-2pm - ECHA 2 150

The screenshot shows the University of Alberta website. At the top, there is a navigation bar with links for 'My UAlberta', 'Find a Person', 'ONEcard', 'Bear Tracks', 'Maps', 'Email & Apps', 'eClass', and 'Libraries'. A search bar is also present. Below the navigation bar is a green header with the University of Alberta logo and the text 'UNIVERSITY OF ALBERTA'. To the right of the logo are links for 'Prospective Students', 'Current Students', 'Faculty & Staff', and 'Alumni & Friends'. Below the header is a horizontal menu with links for 'WHY UALBERTA?', 'FACULTIES & PROGRAMS', 'RESEARCH', 'ADMISSIONS', 'CAMPUS LIFE', and 'NEWS & EVENTS'. The main content area features a large banner with a background of blue and white data points. A white box in the center of the banner contains the text: 'Research Data Management Week', 'April 28 - May 2', 'Manage, preserve, protect and share research data.', and a green button labeled 'FIND OUT MORE'. Below the banner is a news article titled '\$348.5 Billion Engine' with a yellow button labeled 'Apply Now' and two other buttons labeled 'Careers' and 'Give to the U of A'. The article text reads: 'Landmark study by Faculty of Business professors Tony Briggs and Jennifer Jennings highlights the economic and social impact of UAlberta alumni. See the results.'

Support for data stewardship

Recent examples:

- KIAS CRAFT grant
- Data purchase program
- Canadian Polar Data Network (CPDN), Canadian High Arctic Research Station (CHARS)
- Ice Core Archive
- Cloud to local preservation, e.g. genomics, civil engineering
- Preservation & stewardship
 - Canadian Writing Research Collaboratory



Library Human Resources

- RDM identified as top 3 priority of the Libraries
- Digital Initiatives unit - metadata, repositories, digital preservation, data curation, spatial data, HuCo, digitization, etc.
- raising skill level of all librarians
 - awareness and advocacy
 - training
 - direct support, e.g. DMPs, metadata, archiving, sharing

Building capacity - expertise and knowledge

CLIR/DLF Postdoctoral Fellows in Data Curation

Health &
human subject data
(2013-2015)

Web archiving &
digital scholarship
(2014-2016)



Towards a better understanding of value and cost

4C Collaboration to Clarify the Costs of Curation

Search...

Login Register Site Map SHHKT

Home About 4C Work Packages Community Resources Roadmap News and Comment

Community Resources Outputs and Deliverables D3.1--Summary of Cost Models

In This Section

INVESTING IN OPPORTUNITY: CONFERENCE

OUTPUTS AND DELIVERABLES

THE DIGITAL CURATION SUSTAINABILITY MODEL

D1.3--SUSTAINABILITY & BENEFITS REGULATION PLAN

D2.1--BASELINE STUDY OF STAKEHOLDER & STAKEHOLDER INITIATIVES

D2.2--STAKEHOLDER REGISTRY

D2.3--FINAL STAKEHOLDER REPORT

D2.4--FINAL REPORT ON OUTREACH EVENTS

D2.5--PROJECT COMMUNICATION PLAN

D2.6--REPORT ON COMMUNICATIONS ACTIVITIES

D2.8--CURATION COSTS EXCHANGE

D3.1--EVALUATION OF COST MODELS AND NEEDS & GAPS ANALYSIS

D3.1--SUMMARY OF COST MODELS

D3.2--COST CONCEPT MODEL AND GATEWAY SPECIFICATION

D3.3--CURATION COSTS EXCHANGE FRAMEWORK

D4.1--A PRIORITISED ASSESSMENT OF THE INDIRECT ECONOMIC DETERMINANTS OF DIGITAL CURATION

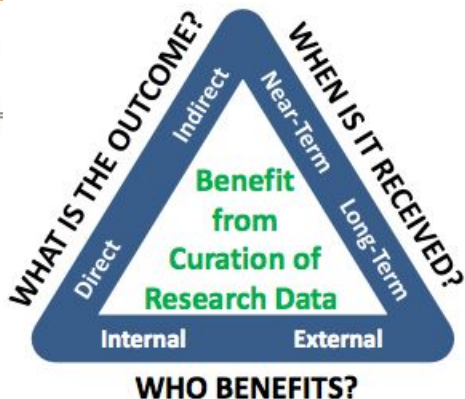
T4.2--DRAFT ECONOMIC

Summary of Cost Models

Here you see a list of ten current and emerging cost and benefit models that the 4C project has evaluated. If you follow the links you will find a short description of each of the models and a presentation of their core features structured in tables to allow for easy comparison.

The tables present among other things what information assets the models can handle, which activities they cover, which cost elements and variables they account for, and references for further reading. The intention of these summaries is to provide potential users with a quick overview of what models they may want to use or have a closer look at.

ID	Name	Acronym	Owner
1	Test bed Cost Model for Digital Preservation	T-CMDP	National Archives of the Netherlands
2	NASA Cost Estimation Tool	NASA-CET	National Aeronautics & Space Administration (NASA)
3	LIFE³ Costing Model	LIFE3	University College London and the British Library
4	Keeping Research Data Safe	KRDS	Charles Beagrie Ltd
5	Cost Model for Digital Archiving	CMDA	Data Archive and Networking Services (DANS)
6	Cost Model for Digital Preservation	CMDP	Danish National Archives and the Danish Royal Library
7	DIP4LIB Cost Model	DIP4LIB	The German National Library
8	PrestoPRIME Cost model for Digital Storage	PP-CMDS	The PrestoPRIME Project
9	Total Cost of Preservation	CDL-TCP	California Digital Library
10	Economic Model of Long-Term Storage	EMLTS	David Rosenthal



LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Enter search term

For Students For Staff For Alumni For Press

STUDY FACULTIES RESEARCH ABOUT US NEWS & EVENTS

Home -- Library & Archives -- Specialist Services -- Research Data Management -- (Produce a DM Plan)

Calculating Data Management Costs

As a general rule, the cost of applying good data management practice is lower if it is performed on an ongoing basis, rather than applied retroactively. Funded researchers are expected to plan their approach to data capture, management and sharing at the earliest opportunity.

The School recognises that the creation, management and sharing of research data carries a cost, requiring the allocation of staff time and other resources. It encourages researchers to consider the resource implications of their research project and write these costs into research proposals whenever possible. Consult the LSHTM RDM Policy for further details.

Funder's position on Data Management costs

Many research funders, including the Medical Research Council, Wellcome Trust and Gates Foundation, encourage applicants to include data management and sharing costs in research bids. Consult data management advice provided by your specific funder for details.

RCUK guidance states data management and sharing costs may be claimed, with a small number of caveats:

- The Data Management Plan should describe activities that incur costs and provide justification the allocation of resources
- No expenditure can be 'double funded', i.e. a service that is centrally supported by indirect costs must not be included as a direct cost as well
- The budget and Justification of Resources section should broadly indicate where RDM costs will be incurred, where possible. E.g. data capture and cleaning, data curation and preservation, data sharing
- Post-project costs may be claimed as a direct cost if it is invoiced before the project end, e.g. PDFS (Pay Once, Store Forever)
- Long-term storage should not be included if data is expected to be deposited in a research council funded repository
- Unforeseen data management costs incurred after project end will not be supported by the funder

Resources

General Information

Guidance

Using the Archives

Specialist Services

- Open Access
- Research Data Management
 - Introduction to RDM
 - Produce a DM Plan
 - Create and Organise Data
 - Keeping Data Securely
 - Documenting your Data
 - Create and Preserve Data
 - Share Data
 - Cite Data
 - Deposit Data
 - Freedom of information
- Data Protection
- Records Management
- Bibliometrics

Diagram: Keeping Research Data Safe Factsheet Cost issues in digital preservation of research data (2011) http://www.beagrie.com/static/resource/KRDS_Factsheet_0711.pdf



UNIVERSITY OF ALBERTA
LIBRARIES

library.ualberta.ca/researchdata

Geoff Harder
Associate University Librarian
Information Resources and Digital Initiatives
geoffrey.harder@ualberta.ca