Research Data Management Faculty Practices:
A Canadian Perspective

Cristina Sewerin
Dylanne Dearborn
Angela Henshilwood
Michelle Spence
Tracy Zahradnik

July 6, 2015

IATUL Conference 2015
Hannover, Germany
Outline

National & institutional contexts

Goals & methods

Results, discussion

Conclusions, next steps
Capitalizing on Big Data: Agencies launch consultation on digital scholarship

Canada’s federal research granting agencies—the Social Sciences and Humanities Research Council (SSHRC), the Natural Sciences and Engineering Research Council (NSERC), the Canadian Institutes of Health Research (CIHR), as well as the Canada Foundation for Innovation (CFI), and in collaboration with Genome Canada—have joined forces to help address digital infrastructure challenges through the consultation document Toward a Policy Framework for Advancing Digital Scholarship in Canada.

Canada stands at the edge of an information and communications revolution. Dynamic new technologies are enabling, accelerating and influencing deep conceptual changes in the research environment, the economy and society.

Many elements of the digital infrastructure required to facilitate advanced research in this new environment are now in place. At the same time, the potential of data-intensive research is quickly exceeding our ability to manage and to grow the digital ecosystem required to meet 21st century research needs, and thus contribute to Canada’s social and economic prosperity.

To help address this challenge, the Agencies in the consultation document are proposing a collective realignment of funding policies that will help promote excellence in data-management practices in
Goals of research survey

• determine how U of T science and engineering faculty and postdoctoral fellows manage and share research data beyond their project

• determine how University of Toronto Libraries (UTL) might help to facilitate data management activities

• understand some of the differences in research data management practices and needs across disciplines and sub-disciplines
Methods

• Online survey
• Evolution engineering
  + computer science, mathematics, statistics, astronomy, astrophysics, physics, chemistry, earth sciences
• Question development
Results and discussion
Demographics

Population (estimated) 1068

Respondents 95

Response rate 8.9 %

<table>
<thead>
<tr>
<th>Respondent Ranks</th>
<th>Postdoctoral Fellow</th>
<th>Lecturer</th>
<th>Professor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>349</td>
<td>63</td>
<td>704</td>
</tr>
<tr>
<td>Postdoctoral Fellow</td>
<td>19</td>
<td>8</td>
<td>68</td>
</tr>
</tbody>
</table>
Research projects, storage use

![Bar chart showing the number of respondents for different storage volumes and project types.](chart.png)
<table>
<thead>
<tr>
<th>Faculty or Department</th>
<th>Mentioned Repository</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faculty of Applied Sciences &amp;</td>
<td>American Concrete Institute, CRAPome, European Bioinformatics Institute (EBI), GitHub, National Academy of Sciences, NIST: X-Ray Photoelectron Spectroscopy (XPS) Database, NEEShub, ONE-ITS, Open Data - Toronto, PhysioData Project: The Virtual Physiological Human, ProteomeXchange, other institutional repositories</td>
</tr>
<tr>
<td>Engineering</td>
<td></td>
</tr>
<tr>
<td>Department of Astronomy &amp; Astrophysics</td>
<td>Canadian Astronomy Data Centre (CADC), Centre de Données astronomiques de Strasbourg/Strasbourg Astronomical Data Centre (CDS), European Southern Observatory, GitHub, Infrared Processing and Analysis Centre (IPAC), Mikulski Archive for Space Telescopes (MAST), SAO/NASA Astrophysics Data System (ADS), other observatories</td>
</tr>
<tr>
<td>Department of Chemistry</td>
<td>arXiv, Cambridge Crystallographic Data Centre (CCDC)/Cambridge Structural Database (CSD), FLUXNET</td>
</tr>
<tr>
<td>Department of Computer Science</td>
<td>Evaluations and Languages resources Distribution Agency, GitHub, Linguistic Data Consortium</td>
</tr>
<tr>
<td>Department of Mathematics</td>
<td>arXiv, HAL</td>
</tr>
<tr>
<td>Department of Physics</td>
<td>academia.edu, AERONET, Canadian Astronomy Data Centre (CADC), Canadian Network for Future Detection of Atmospheric Change (CANDAC), Earth System Grid Federation, GitHub, HopData, ICSU World Data System, Inorganic Crystal Structure Database, NASA Distributed Active Archive Centers*, NASA Langley Research Centre Multimedia Repository, Network for the Detection of Atmospheric Composition Changes (NDACC), Open Crystallography Database, Particle Data Group, Polar Data Catalog, Research Gate, Total Carbon Column Observing Network (TCCON), World Data Center for Paleoclimatology, World Data Center for Atmospheric Trace Gases, World Ozone and Ultraviolet Radiation Data Center (WUDC)</td>
</tr>
</tbody>
</table>

*DAACs include: Alaska Satellite Facility, Atmospheric Science Data Center (ASDC), Crustal Dynamics Data Information System (CDDIS), Global Hydrology Resource Center, Goddard Earth Sciences Data and Information Services Center (GES DISC), Land Processes, Level 1 and Atmosphere Archive and Distribution System (LAADS), National Snow and Ice Data Center (NSIDC), Oak Ridge National Laboratory, Ocean Biology, Physical Oceanography, Socioeconomic Data and Applications Data Center (SEDAC)
RDM support required

- 47.2%: I would be able to draft a data management plan that would address these types of questions without assistance.
- 37.4%: I would be able to draft a data management plan that would address these types of questions, but would prefer to have assistance and/or guided documentation to ensure the success of my application.
- 15.4%: I would need assistance and/or guided documentation to appropriately address some or all of the sections.
Sharing practices

Sharing Practices - Current and Future

- Not currently sharing/Not planning to share
  - Currently using - FASE respondents (n=44)
  - Currently using - Other respondents (n=95)
  - Consider using in the future - FASE respondents (n=44)
  - Consider using in the future - Other respondents (n=94)

- Share by personal request only
  - Currently using - FASE respondents (n=44)
  - Currently using - Other respondents (n=95)
  - Consider using in the future - FASE respondents (n=44)
  - Consider using in the future - Other respondents (n=94)

- Share online with restricted access
  - Currently using - FASE respondents (n=44)
  - Currently using - Other respondents (n=95)
  - Consider using in the future - FASE respondents (n=44)
  - Consider using in the future - Other respondents (n=94)

- Upload online to an institutional or personal website
  - Currently using - FASE respondents (n=44)
  - Currently using - Other respondents (n=95)
  - Consider using in the future - FASE respondents (n=44)
  - Consider using in the future - Other respondents (n=94)

- Include as part of supplementary material files to a journal publisher
  - Currently using - FASE respondents (n=44)
  - Currently using - Other respondents (n=95)
  - Consider using in the future - FASE respondents (n=44)
  - Consider using in the future - Other respondents (n=94)

- Deposit in a general or discipline-specific repository
  - Currently using - FASE respondents (n=44)
  - Currently using - Other respondents (n=95)
  - Consider using in the future - FASE respondents (n=44)
  - Consider using in the future - Other respondents (n=94)
Storage
Cross institutional sharing
Conclusions, next steps

• more detailed analysis
• cross institutional sharing
• strong response: need or want DMP support services, institutional repository
• wait to see what actually happens with funding mandates
Danke schön!
Thank you!
Questions?
Cristina Sewerin
cris.sewerin@utoronto.ca
Photo credits

- news.utoronto.ca
- you.ubc.ca
- www.queensu.ca
- en.wikipedia.org