DRAFT TRI-AGENCY STATEMENT OF PRINCIPLES ON DIGITAL DATA MANAGEMENT *

Michael Miller, AVPR
August 27, 2015

PREAMBLE

• **Research data** – recorded material that validates research findings and results, and enables reuse or replication

• The Tri-Agencies (CIHR, NSERC, SSHRC) are strong advocates for making the results of the research they fund as accessible as possible.
Related Initiatives

• OECD Declaration on Access to Research Data from Public Funding (2004),

• Open Government Declaration (2011),

• G8 Science Ministers Statement (2013),

• Government of Canada’s Action Plan on Open Government (2014) aims to maximize access to the results of federally-funded research to encourage greater collaboration and engagement with the scientific community, the private sector, and the public.
• The Agencies believe that research data collected with the use of public funds belong to the fullest extent possible in the **public domain**, available for reuse by others, and they strongly support the creation of a robust and **efficient environment for data stewardship** in Canada and internationally.

• They have encouraged data stewardship through
  – SSHRC’s *Research Data Archiving Policy* (1990),
  – *Tri-AgencyFramework: Responsible Conduct of Research* (2011), and
Expectations

• Data management planning

• Constraints and obligations
  – Commercial. Legal and ethical obligations

• Adherence to evolving standards and best practices

• Collection and standards
  – Software and formats that ensure secure storage and access well beyond the duration of the project
• **Metadata**
  
  – All research data should be accompanied by metadata that accord with community best practice to enable future users to access, understand and reuse the data.
  
  – Metadata generally state who created the data and when, and may also include information on how the data were created, their quality, accuracy and precision, as well as other features necessary to facilitate understanding and reuse.
• Preservation, Retention and Sharing
  – All research data resulting from agency funding should normally be **preserved in a publicly accessible, secure and curated repository** or other platform for discovery and reuse by others.

  – To determine whether data should be shared and preserved, researchers should consider the data needed to **validate research findings and results**, support **replication and reuse**, and consider the potential benefit to their own fields of research, fields other than their own and society at large.

  – The **retention period** for data should be determined in accordance with community best practices and relevant policies.
• Timeliness
  – **Data should be shared as early as possible** in the research process when they are considered to be informative and of appropriate quality.

  – Data release can be staged as research progresses, starting with metadata, but **data should be shared no later than upon the publication of results**. Where possible, data should be linked to relevant publications.

  – **A defined period of exclusive use of data for primary research is reasonable in some cases.**
• Acknowledgement and Citation

• Efficient and Cost Effective
  – Data management should be efficient and cost-effective.
  – Not all data may need to be shared or preserved – costs and benefits of doing so should be considered in the data management planning process.
Researcher Responsibilities

• **Incorporating data management best practices into their research, and developing data management plans** to guide the responsible collection, formatting, preservation and sharing of their data throughout the entire lifecycle of a research project and beyond;

• Following the requirements of applicable institutional policies and professional or disciplinary standards;

• **Acknowledge and cite** datasets that contribute to their research;

• Staying abreast of **standards and expectations** of their disciplinary community.
Research Institution Responsibilities

• **Providing their researchers with an environment that enables world class data stewardship practices**, as well as delivering, or supporting access to, repositories or other platforms that securely preserve, curate and provide continued access to research data;

• **Monitoring compliance of researcher data management practices with legal, ethical and commercial considerations** in the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – 2nd edition*, *the Tri-Agency Framework: Responsible Conduct of Research*, and other relevant policies;
• Providing their researchers with **guidance to properly manage their data** including the development of data management plans;

• Recognizing data as an important research output and fostering excellence in data management;

• **Promoting the importance of data management to researchers, staff and students**;

• **Developing their own data management policies** and ensuring that these policies and can **accommodate the rapidly evolving research communities’ best practices**.
Research Communities Responsibilities

• Developing data management standards, or promoting existing standards, and working collaboratively to review and improve these standards;

• Recognizing data as an important research output and fostering excellence in data management within their research community;

• Identifying and encouraging the use of specific repositories and platforms.
Research Funder Responsibilities

• **Developing policies and requirements** that facilitate and recognize responsible data management;

• **Providing applicants with clear information and guidance** with regard to fulfilling data management requirements;

• Recognizing data as an important research output;

• **Promoting the importance of excellent data management**;

• Where appropriate, **providing peer reviewers with guidance and developing assessment material for including data management considerations in the application assessment process**.
Research Computing Advisory Committee

• Review proposed institutional responsibilities
  – environment for data management
  – guidelines and best practices
  – monitoring compliance