



Building Research Data Management Infrastructure in Canada from the Bottom-up



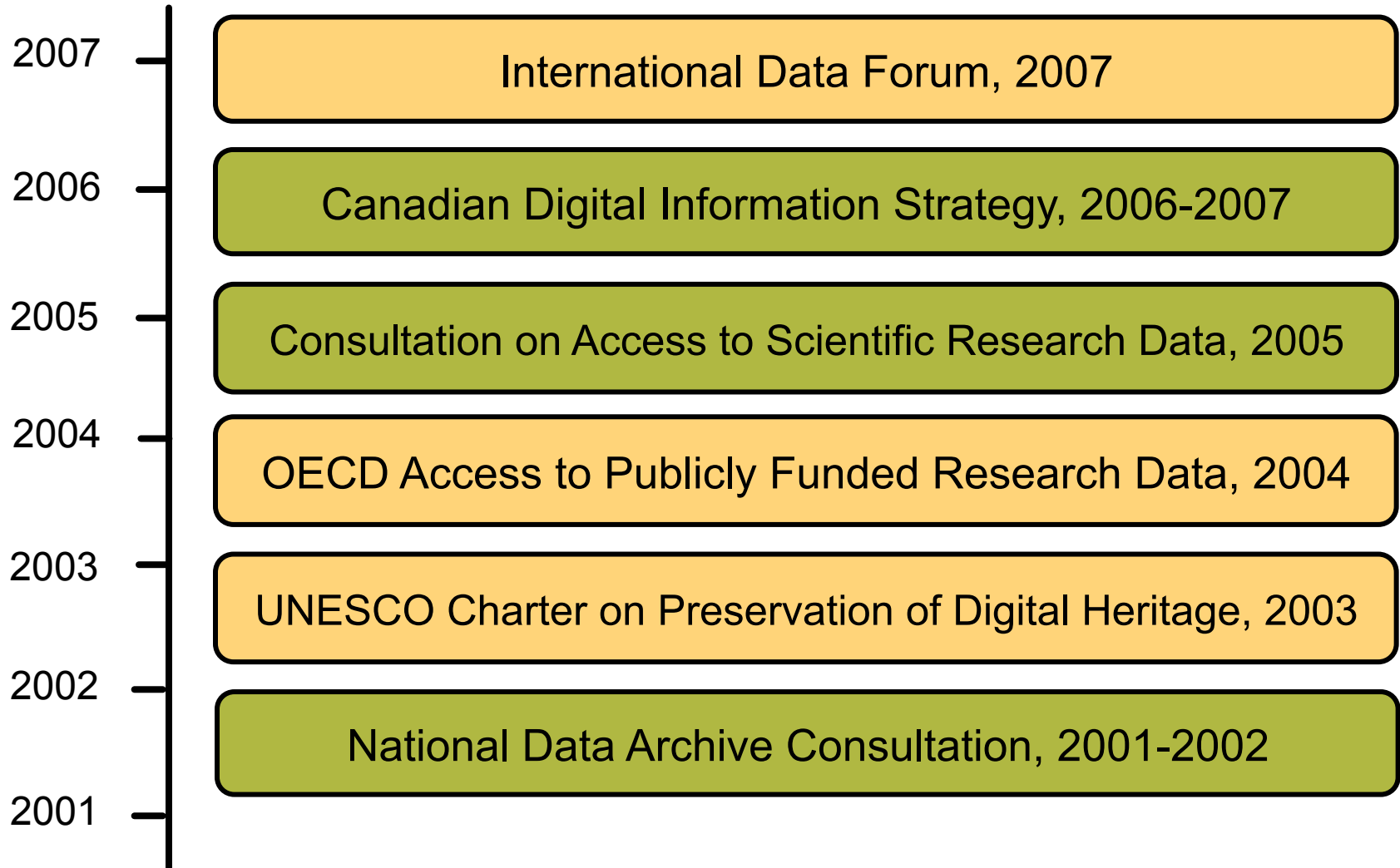
Charles (Chuck) Humphrey
University of Alberta

Where's Canada today?

The following three statements best summarize the situation in Canada today around research data management and preservation.

1. A strategic shift has occurred over the past decade from building a national data preservation institution to building national research data management infrastructure.
2. Building this national research data infrastructure is taking place from the bottom-up.
3. Building this infrastructure from the bottom-up requires intentional, collaborative actions. The driving principle is one of cooperation, not control.

1. Shift from institution to infrastructure



Understanding infrastructure

[C]yberinfrastructure is the set of organizational practices, technical infrastructure, and social norms that collectively provide for the smooth operation of scientific work at a distance (p6).

Understanding Infrastructure: Dynamics, Tensions, and Design

P. Edwards, S. Jackson, G. Bowker

and C. Knobel

January 2007

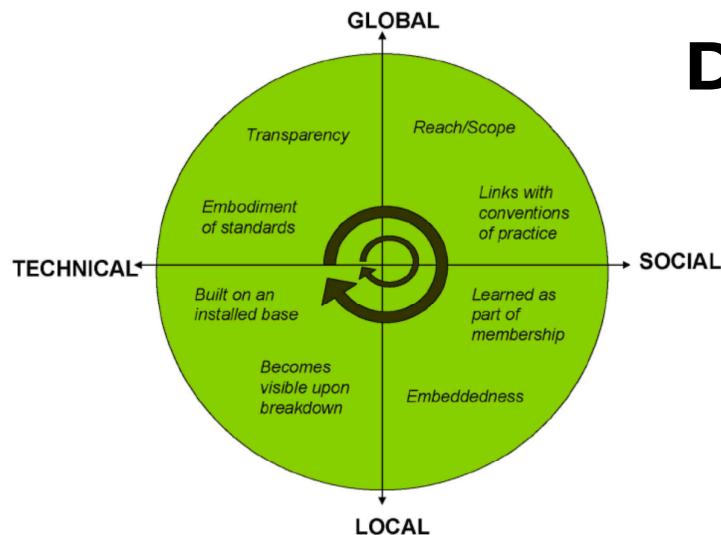


Figure 1. Cyberinfrastructure as distributions along technical/social & global/local axes (diagram courtesy of Florence Millerand).

Research data management infrastructure

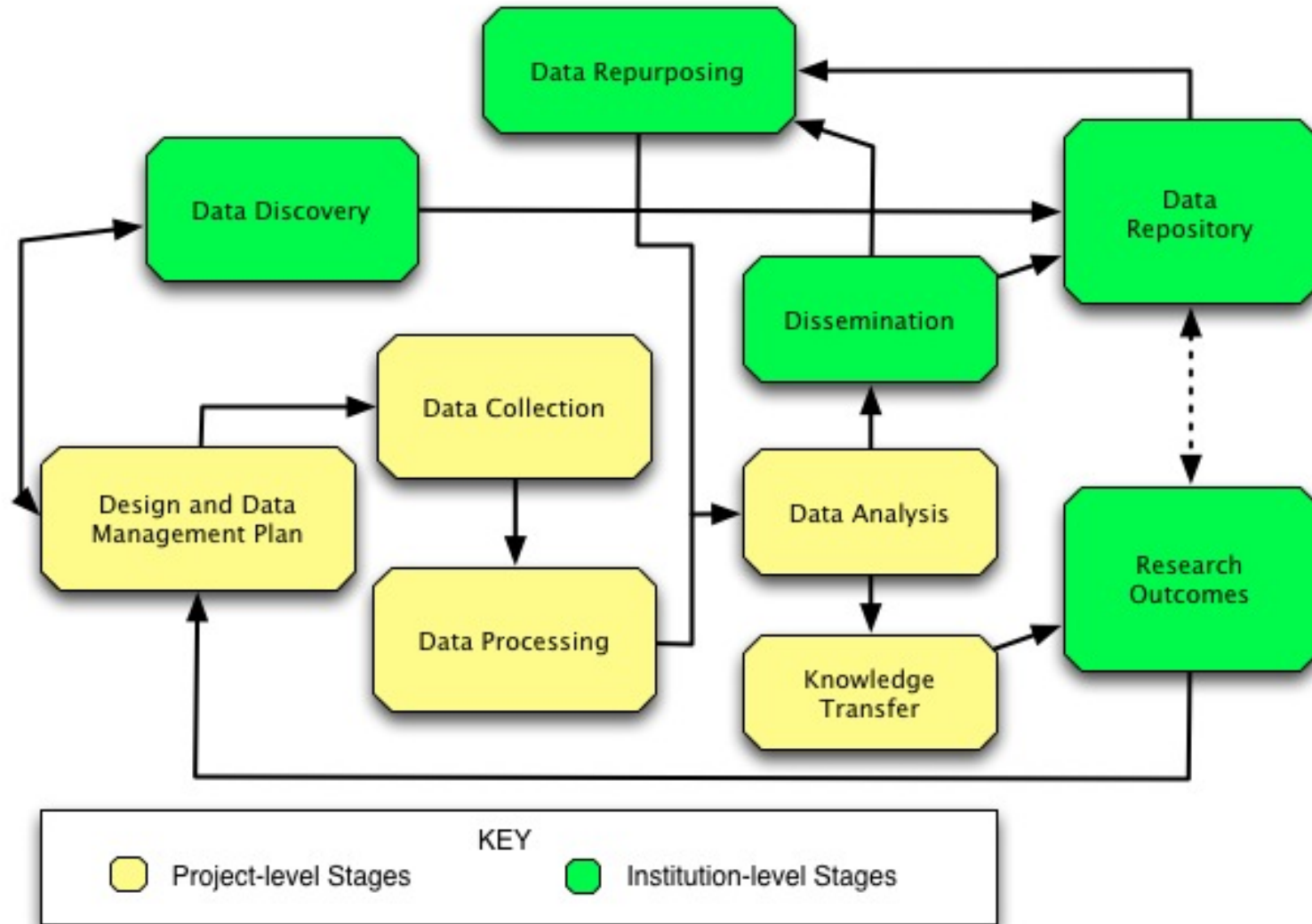
- ▶ RDMI is the configuration of staff, services, and tools assembled to support data management across the research lifecycle and more specifically to provide comprehensive coverage of the stages making up the data lifecycle. It can be organized locally and/or globally to support research data activities across the research lifecycle.

Capitalizing on Big Data: Toward a Policy
Framework for Advancing Digital
Scholarship in Canada
Appendix 4: Definitions

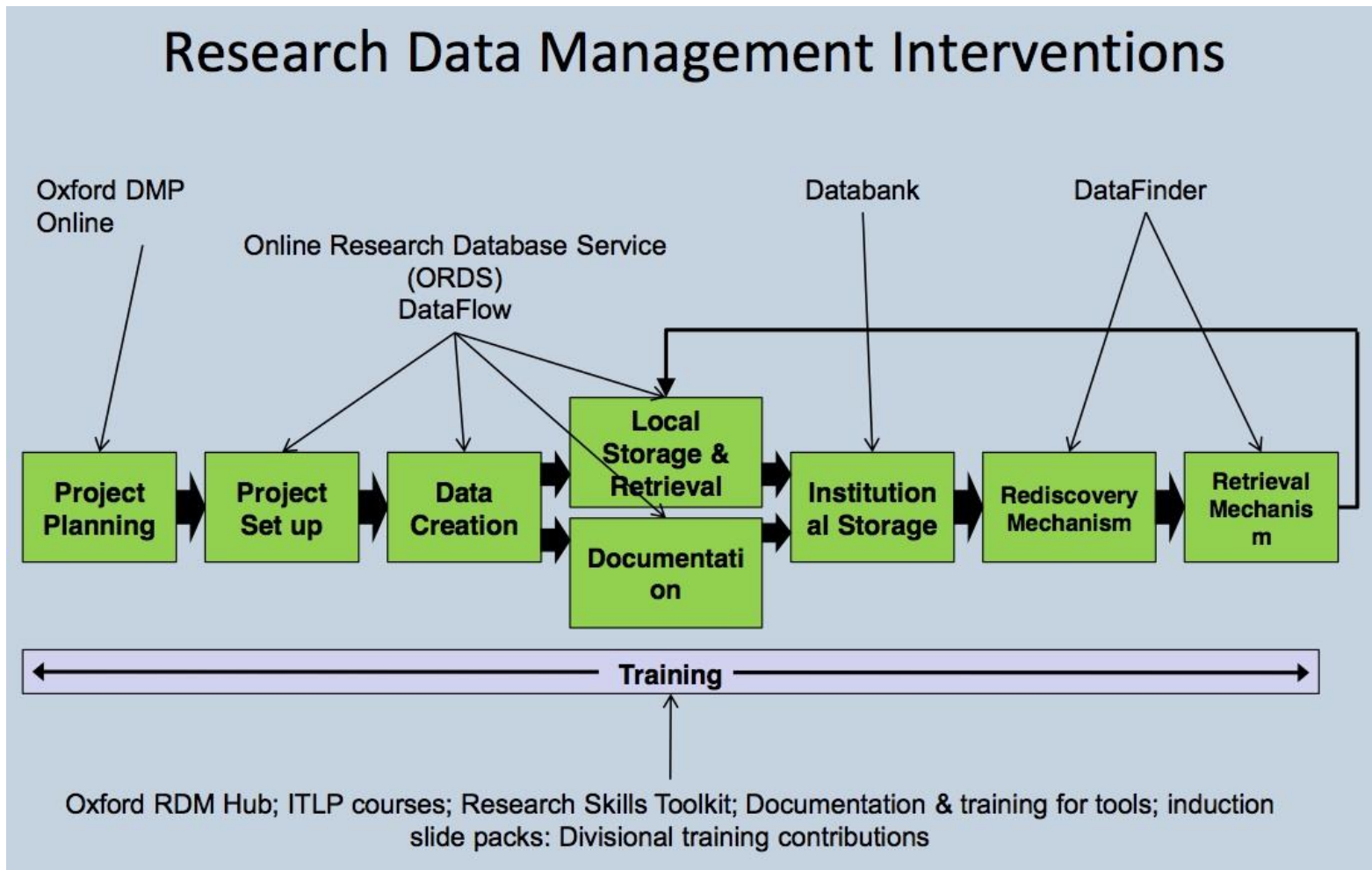
2. Bottom-up development

- ▶ The Brewster Kale principle: just build it
 - ▶ November 12, 2009 CARL Directors' meeting in Ottawa
- ▶ Levels of data stewardship responsibilities
 - ▶ The research project level
 - ▶ The local institutional level
 - ▶ The wider stakeholder level
 - ▶ Across regions, Canada, and the globe
 - ▶ Across domains and research programs
 - ▶ Across sectors

Stewardship levels and the research lifecycle



Institutional Research Lifecycle



Paul Jefferys. **Data Management at Oxford**. March 2012.

Policy at the institutional level

UAPPOL

▼ Policies and Procedures

A - Z Listing

Academic Policies

Advancement Policies

Research Policies

Human Resources Policies

Finance Policies

Administration & Operations Policies

University Relations Policies

Health & Safety Policies

Information Management & Information Technology Policies

Facilities, Properties & Equipment Policies

About UAPPOL

Contact Us



Policy Development
Toolkit

Home > Policies and Procedures > Policy detail

Approval Date: 06/03/2013 **Effective Date:**

Research Policy

Description:

Research Policy

Office of Accountability:

Vice-President (Research)

Approvers:

Board Learning and Discovery Committee (BLDC)

Office of Administrative Responsibility: Vice-President (Research)

Scope:



View Full Policy

PUBLISHED PROCEDURES OF THIS POLICY

- Research Administration Procedure (Roles and Responsibilities)
- Research Records Stewardship Guidance Procedure
- Research Records Stewardship Guidance Procedure Appendix A: Research Records Management and Preservation Guidelines
- Research Records Stewardship Guidance Procedure Appendix B: Research Records Classification Guidelines

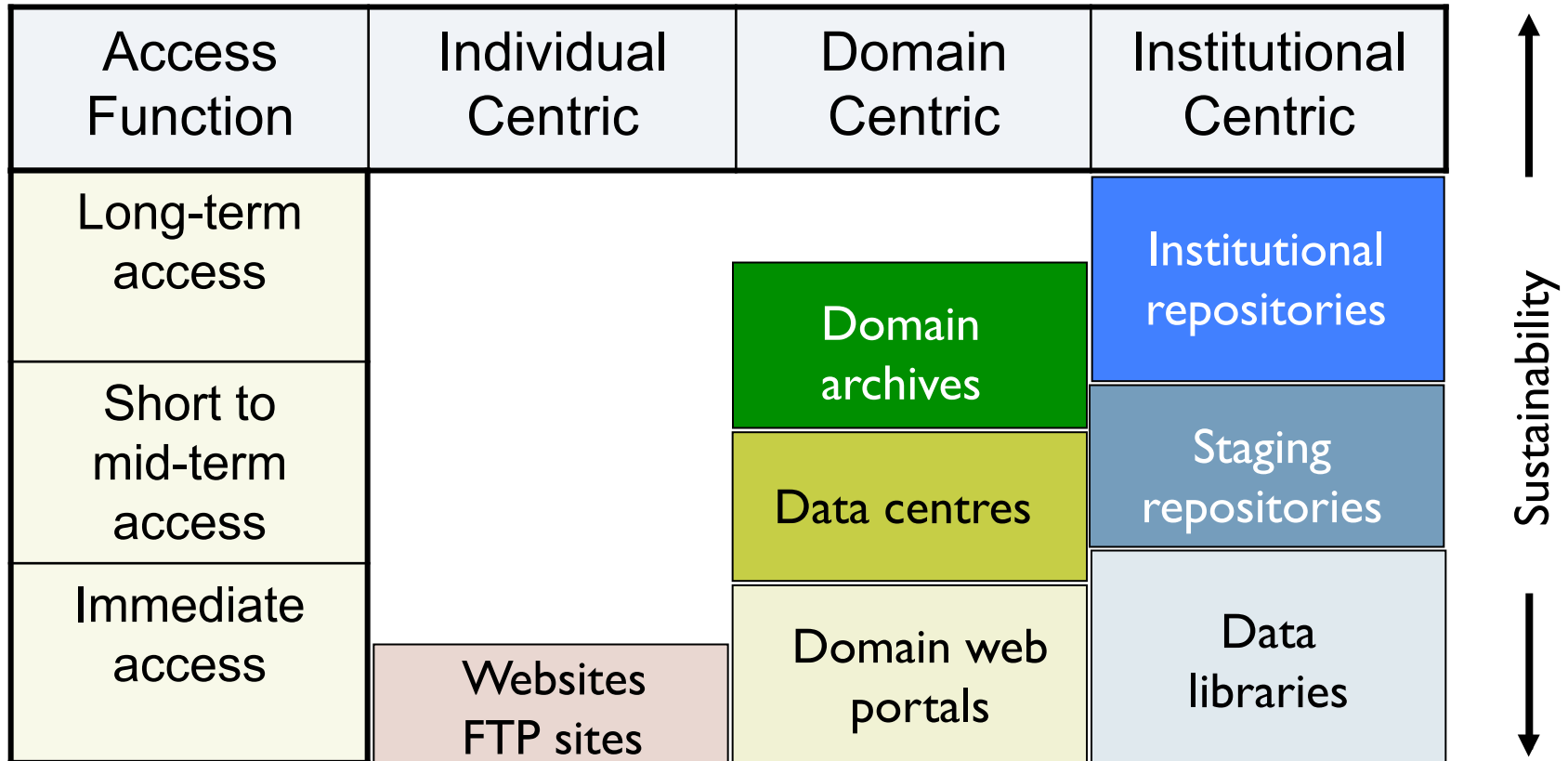
At the wider research stakeholder level

- ▶ A few examples:
 - ▶ Individual institution across sectors
 - ▶ Canada's International Polar Year and the development of the IPY Data Assembly Centre Network and its transformation into the Canadian Polar Data Network
 - ▶ Consortia within region
 - ▶ OCUL/SP cloud storage project
 - ▶ OCUL/SP Dataverse Network
 - ▶ Regional consortia
 - ▶ The Canadian Social Science Research Data Private LOCKSS Network
 - ▶ Shared functionality enhancements to Archivematica for research data
 - ▶ National membership
 - ▶ The CARL Research Data Management Institute

3. Successful bottom-up characteristics

- ▶ Capitalize on the energy driving the sense of urgency around sharing and preserving research data, which is resulting in potential partners across sectors and institutions.
- ▶ As we identify potential partners, we have begun to change the metaphors that we use to describe the organizational representations of research data infrastructure. We have gone from “data landscapes” to “data ecosystem.”

The data landscape



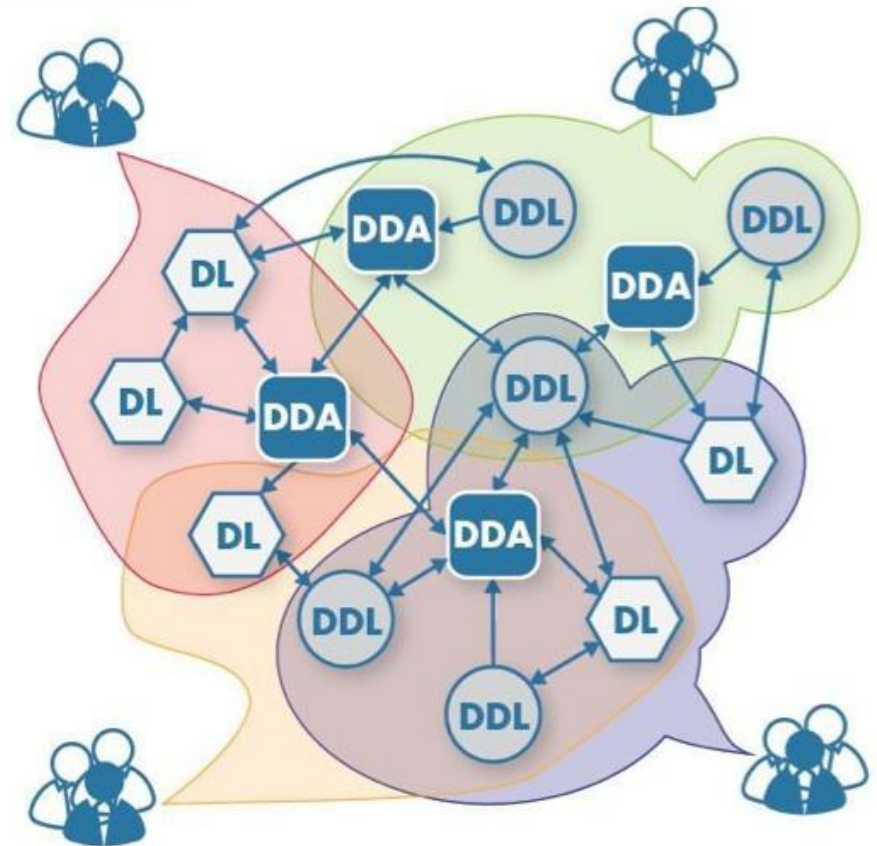
Research data ecosystem



Legend



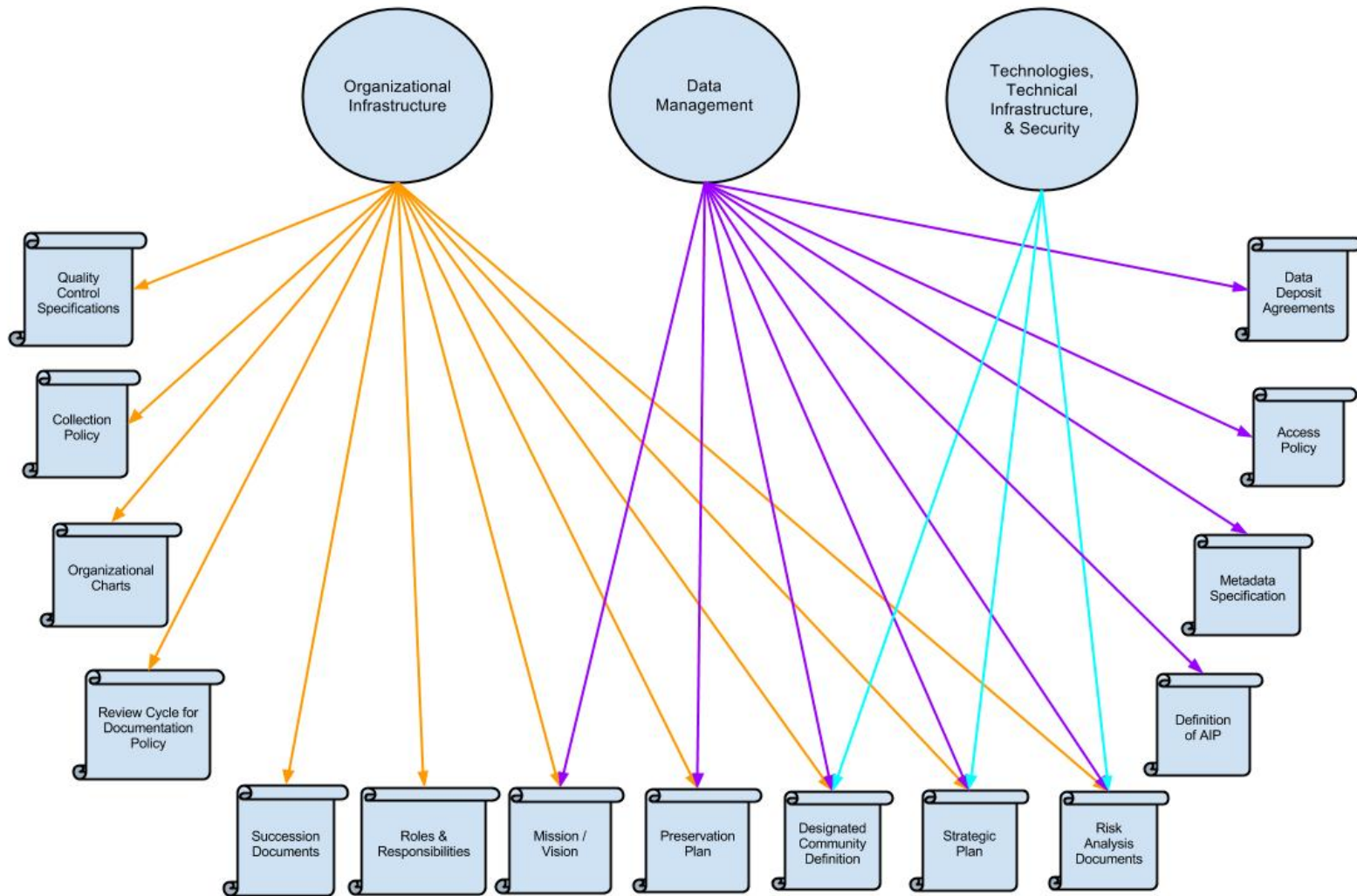
Figure 1 – GRDI2020 Digital Science Ecosystem



Building of a successful collaboration

- ▶ Take steps to build trust among partners, which doesn't always come from the tops of organizations. Let those passionate within organizations find solutions, working together with their counterparts across organizations.
- ▶ Prepare a charter that expresses the norms for working together and the common commitment to the task of research data preservation. A good example is the Charter of the Canadian Polar Data Network (see http://polardatanetwork.ca/wp-content/uploads/CPDN_Governance.pdf)
- ▶ Develop a set of policies to serve as a foundation for the shared research data management infrastructure.

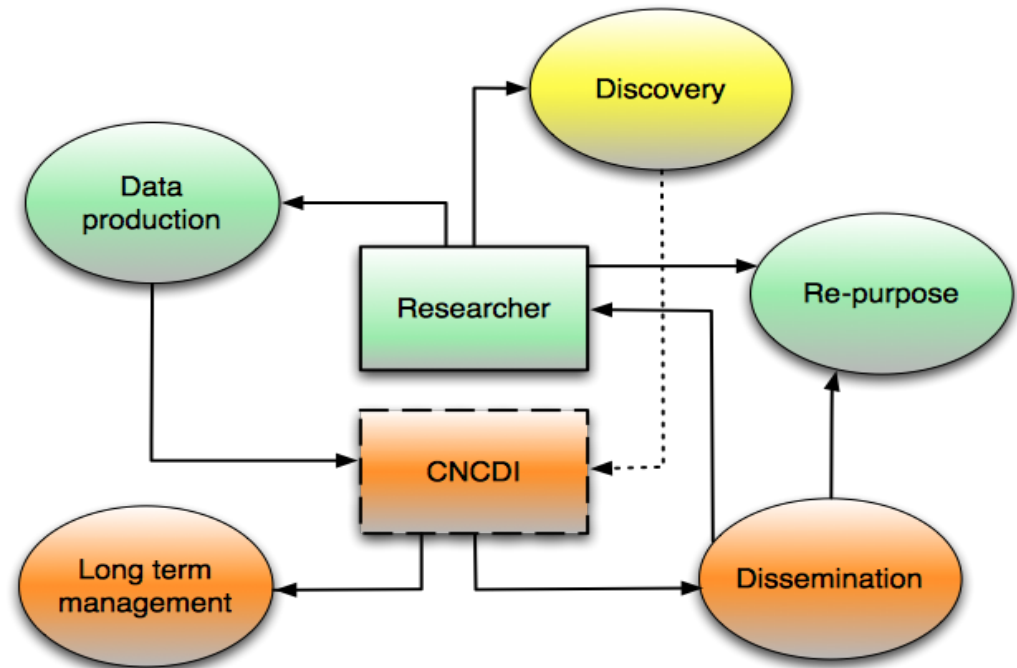
Data policy document framework



Building of a successful collaboration

- ▶ Develop blueprints for new research data management infrastructure in teams across institutions and do what is possible now while laying the groundwork for what can be incorporated in the future.

The CARL Canadian National Collaborative Data Infrastructure proposal as a blueprint.



Building of a successful collaboration

- ▶ Pool resources to get infrastructure in place.
 - ▶ The CWAP project is an example of this approach. This is a jointly funded initiative between the University of Alberta, UBC, and SFU to add research data preservation functionality to Archivematic, a tool developed by Artefactual Inc in Vancouver that produces high quality archival information packages. More recently, OCUL/SP has expressed interest in contributing to this develop and the University of Saskatchewan has a project to extend functionality between Islandora and Archivematica.

Building of a successful collaboration

- ▶ Integrate and be open to new partners, including a variety of designated user communities.

Responsibilities in the Research Environment

- Research Data Infrastructure: Preservation Backbone Network with Staging Repositories integrated within research projects to support data management and Ingest Centres supporting retrospective deposit of data from research. Access to provided through existing or new data centres, data libraries or data warehouses.
- High Performance Computing: Compute Canada
- High Speed Research Network: CANARIE and Can. Access Fed.

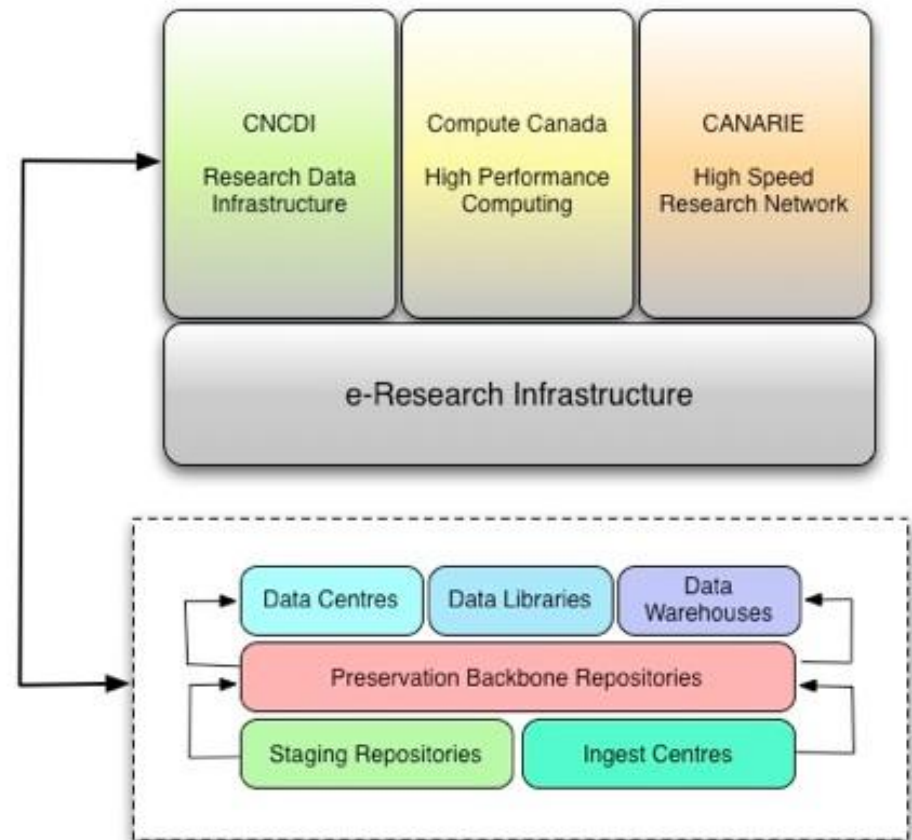


Figure 2. Responsibilities in the research environment

“By what authority ...”

- ▶ Build trust among the communities being served.
- ▶ Demonstrate competencies in delivering services.
- ▶ Develop a positive reputation around trust and competencies.
- ▶ Operate from a data culture that incorporates norms of best practice and in which rewards are only part of the reason for engaging.

An RDMI agenda for the Canadian Research Data Management Network

- ▶ **RDM policy and resource coordination**
 - ▶ Develop, promote, interpret, and review RDM policies
 - ▶ Collaboratively raise resources for joint RDM projects
- ▶ **Services**
 - ▶ Coordinate service delivery across the data lifecycle: planning, managing, sharing, discovering, repurposing, and preserving research data
- ▶ **Tools and technology**
 - ▶ Identify, evaluate, and develop tools supporting DM across the research lifecycle
 - ▶ Identify, evaluate, and develop preservation tools and technology
- ▶ **Expertise**
 - ▶ Upgrade and train DM skills for stakeholders across the lifecycle
 - ▶ Develop and advance RDM specializations
- ▶ **Local or globally**