INNOVATION.CA
CANADA FOUNDATION FOR INNOVATION | FONDATION CANADIENNE POUR L'INNOVATION

CFI's Cyberinfrastructure Initiative

Information Sessions
Draft Call for Proposals

January 2015

Research builds communities
La recherche au service des collectivités

Today's agenda:

Part 1: CFI information session

- Cyberinfrastructure Initiative overview
- A few questions and a few answers
- Other questions: discussion period

Part 2: Compute Canada consultation

Context

Two observations:

- In Canada and internationally, the capacity to generate research data and information has far outstripped our capacity to analyze and mine the data to glean new knowledge or insights.

- This creates significant pressure on Canada's existing advanced research computing infrastructure because it may have reached the limits of its current capacity and its computational services may not be sufficiently aligned to the needs of tomorrow's leading edge research.

Context

Two opportunities:

- Unlocking the full potential of research data by organizing, structuring and integrating data sets and developing analytical tools to mine the data; and

- Upgrade and modernize Canada's advanced research computing platform in a manner that support computationally challenging research and data intensive research.

Key Parameters

"*research will drive the investments"*

- Excellence is the key driver. Seeking to enable:
  - globally competitive,
  - data-intensive and
  - computationally challenging …

  …research
- Competitive element
- Maintain commitment to shared resources

INNOVATION.CA

The Cyberinfrastructure Initiative

INNOVATION.CA

Cyberinfrastructure Initiative: Why?

Seeks to achieve a dual objective:

- Renew and upgrade Canada's advanced research computing platform to support tomorrow's leading-edge research, both computationally challenging and data-intensive; and
- Enhance our capacity to analyze and mine research data through the development of research data infrastructures.

Cyberinfrastructure refers to:

- Advanced data storage, data management, data integration, data mining, data visualization and other shared computing and information processing services that are beyond the scope of a single institution to provide.

Cyberinfrastructure Initiative

"Enhance the capacity of Canadian institutions and researchers to conduct leading-edge research by supporting infrastructure needs of computationally and data-intensive research."

Two-pronged approach:

- A limited number of research data infrastructure projects developed in collaboration with Compute Canada
- Upgrading and modernizing the computational and data storage capacities of the pan-Canadian advanced research computing platform, managed by Compute Canada.

Cyberinfrastructure

Following the CFI definition of research infrastructure:

- Common or shared resources for data intensive and computationally challenging research. This includes:
  - the design and development of integrated information databases and analytical tools to exploit these resources;
  - the creation of specialized software tools required to effectively use the data resource;
  - personnel for the development and integration of federated databases, and the development of software tools.
- Intended for resources that require computational and data storage capacity beyond what can be provided by individual institutions

INNOVATION.CA

Cyberinfrastructure Initiative

Overall $50M budget, planning for 2 competitions under each Challenge

CHALLENGE # 1: research data infrastructure projects

- institutions and researchers to come together to form consortia and propose research data infrastructure projects that create tailored, shared and integrated data resources (e.g. databases, data repositories) capable of enabling leading-edge research.
- involve multi-institutional consortia of researchers, data scientists, data analysts and software developers, to design and construct shared research data infrastructure projects for community of researchers to advance their research programs.

Between 5 and 10 projects funded per competition, up to $20 M over two competitions

Cyberinfrastructure Initiative

CHALLENGE # 2: modernize and upgrade the advanced research computing platform

The advanced research computing community, through Compute Canada, to propose a set of capabilities and services that will meet the needs of Canadian researchers conducting data-intensive and computationally challenging research over the next five years.

- **Stage 1 for pressing current needs, Stage 2 for support of projects under Challenge 1**
- **Consultative process, driven by research needs, single proposal with scenarios/options**

At least $30M available for the two Stages

Cyberinfrastructure Initiative: tentative timeline

Key dates	Challenge 1: Research data infrastructure proposals	Challenge 2: Compute Canada
November 2014	1st competition call for proposals released	Call for the Stage 1 proposal to address pressing immediate needs
February 2015	1st competition deadline for submission of expressions of interest	
April 2015		Submission of the Stage 1 proposal
May 2015	1st competition deadline for notices of intent	
June 2015	1st competition invitation to submit full proposal	**CFI DECISION on Stage 1 proposal (tentative)**
October 2015	1st competition deadline for full proposals; 2nd competition call for proposals released	
March 2016	**CFI DECISION on 1st competition proposals (tentative)**	
April 2016	2nd competition deadline for notice of intent	Submission of Stage 2 proposal to support funded research data infrastructure projects and to address continued upgrading and modernizing of the advanced research computing platform
May 2016	2nd competition invitation to submit full proposals	
June 2016		**CFI DECISION on Stage 2 proposal (tentative)**
October 2016	2nd competition deadline for full proposals	
March 2017	**CFI DECISION on 2nd competition proposals (tentative)**	

INNOVATION.CA

A few questions on Challenge 1…

Why is there a suggested upper limit for proposals?

- Looking for projects that are mature and can be operational within 2-3 years;
- These are not general purpose software development initiatives;
- Reason for favouring linkages to existing initiatives in Canada and internationally and that are conducive to adopt, adapt and develop approaches

A few questions on Challenge 1…

We have just begun discussions around forming a consortium. Can we be competitive in the first competition?

- Probably not, but there will be another opportunity in 2016;
- Confident that more support will be forthcoming, provided the experiment shows that it addresses a real need and provides indications of success

A few questions on Challenge 1…

- **How does the CFI define a consortium?**

- Project-, discipline- and domain-specific
- Brings multiple institutions, research teams/groups together (**community of researchers**): critical mass
- Could be international, national, pan-Canadian, regional

A few questions on Challenge 1…

What sources of matching funds are eligible for the research data infrastructure projects?

- The **same** as usual.
- In the case projects linking to international initiatives, **certain costs may be recognized** although the valuation of the contribution can be complex and risky (new contributions, not « recycled » contributions)
- Advisable to speak to the CFI to discuss eligibility and reduce potential exposure.

A few questions on Challenge 1…

Why are computational hardware and data storage equipment not eligible for research data infrastructure projects?

- computational hardware and data storage equipment will not be an eligible expense.
- The CFI expects that the **projects funded under Challenge 1 will require similar capabilities and some common needs, and so…**
- **Economy, efficiency and effectiveness** will be better served by making CC responsible for the development of the technical solution(s) (storage and computation)

A few questions on Challenge 1…

When are we expected to discuss our proposal with Compute Canada? What is expected?

- Compute Canada's resources and efforts would be best utilized in discussions **once the initiative is sufficiently mature and progressed beyond the EOI**.
- CFI will require that **all** those invited to submit full proposals hold formal discussions with Compute Canada to **allow CC to determine if the pan-Canadian platform already has the required capabilities** to support the proposal, or **whether new capabilities will be required**.
- For each full proposal, the **CFI will receive a letter from Compute Canada** outlining the capabilities required to support the proposal and confirming whether their needs can be accommodated through existing capabilities or through new ones to be added (which ones, how much, etc…) .

A few questions on Challenge 1…

What is the likelihood of being successful under Challenge 1 but not under Challenge 2 (Phase 2)?

- The second proposal to be submitted under Challenge 2 is designed to s**pecifically meet the needs** of proposals funded under Challenge 1.
- Seeking the **optimal solution** to support the group of Challenge 1 proposals: merit reviewed

A few questions on Challenge 1…

Do I have to make my data available beyond my immediate collaborators?

- « tailored, shared and integrated data resources »
- expect all consortia members to share data between them; « our data » versus « my data »

The **entire initiative is about shared* and common data** at the national and international scale.

***within (and even beyond) the consortia**

INNOVATION.CA

A few questions on Challenge 1…

What does the CFI intend when it suggests to link with international initiatives?

- This is not a requirement but where appropriate and **consistent with the collective effort of these initiatives** and the benefits that can be derived from such linkages, namely:
  - Reduce duplication;
  - Adopt, adapt and develop approaches;
  - Access to international data resources and integrate Canadian ones with those developed elsewhere.

A few questions on Challenge 1…

Is a proposal to build and develop generic software tools eligible?

- These are **not** general purpose **software development initiatives**.
- The software needs to be **linked to a specific** research data infrastructure project.
- Two or more projects could use the same software tool/interface, in whole or in part.

A few questions on Challenge 1…

How will the CFI address competing or similar proposals?

- Resources are **insufficient to engage in seeding multiple projects**. The CFI is seeking innovative projects and will itself endeavour to be strategic.
- It is up to proponents to conduct their own risk assessment and develop their winning strategy.