Cyberinfrastructure Initiative

Challenge 1

Competition 2 – 2017

Call for Proposals

TABLE OF CONTENTS

1.   PREAMBLE

As part of the 2015 launch of its Cyberinfrastructure Initiative, the Canada Foundation for Innovation (CFI) announced two competitions under Challenge 1 – research data infrastructure. The current document provides information relating to the second competition of Challenge 1 of the Cyberinfrastructure Initiative. It supersedes any information found in the Call for Proposal for the first competition, and should be used when preparing a proposal for the second competition of Challenge 1 of the Cyberinfrastructure Initiative.

2.   CONTEXT

Research is increasingly data intensive

Today, research across all fields of inquiry is increasingly data intensive. Data is an output of research and its analysis often forms the basis for new research hypotheses. As such, it is a powerful enabler of new scientific insights and drives both discovery and innovation. The generation of massive amounts of data by new research capabilities is giving rise to the development of tools, methods and standards necessary to organize and exploit these digital resources. These in turn enable significant economic and social development opportunities in areas such as data management and data analytics. Leading-edge, data-intensive research will only be successful in generating new knowledge if it is supported by robust research data infrastructure tools.

Research data infrastructure

Many of the challenges facing data-intensive research arise from the difficulty of creating tailored, shared and integrated research data infrastructures. Addressing these challenges will require the development of tools, methods and standards to effectively organize, access, mine and analyze massive datasets. Its development requires partnerships within and, in some cases, across research domains[1] so that the shared infrastructure supports a wide range of users. Such partnerships will reduce unnecessary duplication of efforts within and between disciplines. Adopting a collaborative, partnership-based approach requires a collective effort wherein communities of researchers work closely with technical experts (such as data scientists, data analysts and software developers) who are able to build research data infrastructures, and with other stakeholders. In many cases, the expertise required to build research data infrastructures already exists in organizations such as Compute Canada or CANARIE, and should be effectively employed.

The CFI's goal is to support research communities in their efforts to devise optimal ways to organize and use research data resources.

The most promising avenue to address this challenge is to create consortia of researchers and institutions since a single institution is unlikely to possess all of the requisite expertise and resources to do it on its own. Moreover, these collective efforts need not consist solely of developing "made in Canada" solutions. Tapping into existing international initiatives for the development of analytical

---

1 For the purposes of the Cyberinfrastructure Initiative, a research domain is defined as a multi-disciplinary area of research, or a group of areas of research that face common data challenges that can be addressed by the development of tailored, shared and integrated research data infrastructures.

tools, data management methods and standards will ensure that Canadian researchers are in a position to readily access international data resources as well as to integrate Canadian ones with those developed elsewhere and make them available to international collaborators. There may be tools, methods and standards that have been developed internationally that can be adopted in Canada or adapted with minimal effort. Similarly, applicants must consider opportunities to adapt and build on existing Canadian initiatives such as those previously funded by the CFI, CANARIE, the tri-agency or others.

3. THE CHALLENGE

The CFI challenges 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 and data repositories) capable of enabling leading-edge research on significant scientific, social and economic questions. Projects submitted should bring together a **community of researchers from across the country** who share similar challenges linked to the availability of research data. Projects should **address an existing or emerging challenge for this community** through the **development of new tools and applications or novel ways of organizing and using research data** that would enhance the community's capacity to conduct leading-edge research. The entire research data infrastructure component of the initiative is about data-sharing and providing access to any researcher who can exploit and mine the resource to advance knowledge and promote innovation. The projects must be completed within a three-year time frame. Therefore, consortia are encouraged to embrace an "adopt, adapt and develop" approach as well as to link to Canadian and international initiatives whenever appropriate. Other Canadian organizations such as CANARIE and Compute Canada are key stakeholders that could provide insights about the existing landscape and highlight potential synergies with research groups having similar data infrastructure requirements. This will promote efficiency, interoperability and rapid implementation. We also encourage projects that propose to expand and extend existing Canadian data initiatives.

Through this initiative, the CFI expects to support proposals that meet the following three objectives:

- Scientific excellence
- Impact and ongoing relevance
- Feasibility

Once created, these research data infrastructures will be dynamic research resources requiring regular upkeep to ensure that their content is accurate and reliable, and that they remain relevant and able to support leading-edge research efforts over the long term. Consequently, the CFI challenges consortia to develop proposals that demonstrate the sustainability and long-term relevance and usefulness of the research data infrastructure. Some of the ongoing operational funding could be provided through CFI's Infrastructure Operating Fund, but it should not be the sole source.

Consortium

To enable successful design and implementation of the research data infrastructure, it is expected that the pan-Canadian consortium will be composed of both scientific and technical experts. More and more, complex research data projects evolve across disciplines. Thus, cyberinfrastructure projects require a close collaboration between research disciplines (e.g. biologists, chemists, sociologists, etc.) and technical experts at the different stages of the project. For the purpose of the Cyberinfrastructure Initiative, the following definitions apply:

Scientific experts: Scientific experts are subject matter experts who will be involved from the early stage of the project in defining the research questions and the data requirements, based on gaps and opportunities identified by a broader research community. They will both serve as advisors during the development of the research data infrastructure and be the end users once it is fully operational. Consequently, their engagement throughout the project will be critical to the success of the endeavor.

Technical experts: Technical experts are software developers, business analysts, data specialists, etc. who will be involved in the day-to-day development of the research data infrastructure. They are well aware of technical solutions and are not necessarily experts in the scientific area that will benefit from the research data infrastructure.

The complexity of the design and development of the research data infrastructure requires a close collaboration between technical and scientific experts to clearly define the end users from the beginning of the process and to develop a research data infrastructure that address current and emerging needs of the research community.

4.   FUNDS AVAILABLE

The CFI expects to fund between five and 10 proposals. As such, institutions should be very selective about which Notices of Intent they participate in. Given the small number of awards anticipated, an institution cannot submit a proposal for a project funded in the first competition.

Following standard practice, the CFI will contribute up to 40 percent of the eligible costs of an infrastructure project. Each project will be limited to a request of a maximum of $2 million from the CFI.

The CFI recognizes the need to contribute to the operating and maintenance costs of the research data infrastructure projects. Therefore, funded projects under Challenge 1 will receive an Infrastructure Operating Fund (IOF) contribution of 30 percent of the CFI award. However, the ongoing costs associated with the upkeep of the resource will be the responsibility of the consortium members and the administrative institution.

5.   COMPETITION TIMELINES

Key dates	Milestones
March 02, 2017	Draft Call for Proposals released
March 31, 2017	Call for Proposals released
May 12, 2017	Deadline to submit notices of Intent
End of June 2017	Invitation to submit proposal
October 13, 2017	Deadline for proposals
March 2018	CFI funding decision

6.   ELIGIBILITY

Institutional eligibility

Canadian universities, colleges, research hospitals and non-profit research institutions that have been recognized as eligible by the CFI can apply.

Eligible costs

The CFI will contribute to eligible costs up to the point where the research data infrastructure is ready to be used by the larger community of researchers. For added clarity, this includes costs for:

- The design and development of integrated research data resources;
- The creation of specialized software and analytical tools, methods and standards required to effectively exploit the data resources;
- The personnel required to design, develop, build and integrate the data resource, software and analytical tools, methods and standards;
- Personal computing needs of the above-mentioned personnel;

- The acquisition of software and analytical tools;
- Test beds, where they are well justified and cannot be provided by Compute Canada, provided they do not represent a significant proportion of the total project costs;
- The renovation of existing space, for the personnel required to develop and build the research resource.

Ineligible costs

The costs for major computational hardware and data-storage that are beyond the personal computing needs of the personnel directly involved in the development of the research data infrastructure tool are not eligible. Such shared computational and storage resources should take advantage of the pan-Canadian advanced research computing platform managed by Compute Canada. This will promote the efficient, effective and economical use of the computation and data-storage infrastructure of this pan-Canadian platform. When necessary, applicants should therefore engage Compute Canada early in the process to determine how and if its platform can accommodate the needs of the consortium.

The costs associated with primary data collection are also ineligible. Please refer to the CFI Policy and program guide for more information on eligible costs.

Partner contributions

The participation of partners (international members, private sector) is welcome where appropriate, and the CFI could recognize certain costs and contributions associated with their involvement. For example, partner contributions could include the procurement of software and analytical tools and the provision of professional personnel or external experts assigned to the project. The CFI recognizes that the valuation of in-kind contributions from partners could be complex and risky; this is particularly true for projects that are linked to international initiatives. Potential applicants are encouraged to discuss such situations with the CFI early in the process of developing their Notice of Intent or proposal.

As per the established policy, the value of datasets that are not otherwise sold to third parties is considered null. The CFI would only recognize the incremental cost to customize the dataset.

To be eligible for funding, in-kind contributions from external partners and cash expenditures by the institutions must have taken place on or after April 1, 2016. Expenditures are considered incurred when goods are received, services have been rendered or work has been performed.

Access to the research data resources

While it is expected that consortium members represent a critical mass of Canadian researchers in the area, the entire research data infrastructure component of the initiative is about data sharing and providing access to any researcher who can exploit and mine the resource to advance knowledge and promote innovation. Consequently, the research data infrastructure should be made available to all Canadian researchers, regardless of their affiliation.

The CFI acknowledges that some data may be sensitive due to confidentiality or security. Such issues can be addressed in the design of the research data infrastructure, access policies and regimes, etc. The CFI expects to receive applications for a wide range of projects from very open datasets (even public) to projects requiring controlled environments. The CFI will not be prescriptive and will leave it to researchers and their institutions to devise appropriate policies and regimes that maximize user access, in keeping with the particular nature of the proposed research data resources.

Access to data does not equate with free access. Access fees or membership fees can be charged and thereby contribute to the sustainability of the research data infrastructure.

7.   REVIEW AND DECISION MAKING

Expression of interest

Contrary to the first competition, expressions of interest are not required for this competition.

Review of Notices of Intent

Given the available competition budget, the limited number of awards that will be made and the level of effort required to develop proposals, institutions intending to submit research data infrastructure project proposals will be required to submit a Notice of Intent (NOI). The NOIs will be subject to merit review by a Multidisciplinary Assessment Committee (MAC). This merit review will identify those NOIs that best fit the objectives of the Cyberinfrastructure Initiative. The following criteria will be used by the MAC to reach its decision:

- **Scientific excellence:** The proposal relates to a field in which Canada is recognized for having significant research strengths. The research data infrastructure project addresses opportunities or gaps identified by the research community and will enable established and emerging leaders to remain internationally competitive;

- **Impact and ongoing relevance**: A critical mass of Canadian scientific experts is actively involved throughout the project and use of the research data infrastructure will be maximized through optimal access mechanisms; The research enabled has the potential to lead to significant tangible benefits to society, health, the economy and/or the environment;

- **Feasibility:** The research data infrastructure project's scope is clearly defined and the technical experts have the necessary expertise and experience to design, develop and deliver the project within 36 months.

The list of NOIs will be posted on the CFI website. The CFI will invite proposals from the best NOIs that in total represent approximately three times the available competition budget. Given these limitations, the CFI encourages institutions to be very selective about which NOIs it participates in.

Review of proposals

Expert Committee review

In this stage of the merit review, expert committees will be convened to assess the strengths and weaknesses of the proposals in relation to the six assessment criteria. These committees will be tasked with recommending to the MAC those proposals that meet the standard of excellence for the competition and with recommending the amount that should be awarded to each proposal. Proposals not recommended by the Expert Committees will not be considered by the MAC.

Expert Committees will review groups of similar or related proposals. These committees will convene either by teleconference or in person. Additional written external reviews may be solicited and used by the Expert Committees if supplementary expertise is required.

The Expert Committees will evaluate the proposal in accordance with the following criteria:

- **Research:** Once completed, the research data infrastructure will enable research activities that are timely, innovative and at the leading edge internationally.

- **Research data infrastructure:** The research data infrastructure is necessary and appropriate to conduct the proposed research activities and builds, when appropriate, on existing national or international data resources. The scope and requirements of the project are clearly defined and it can be commissioned within 36 months.

- **Scientific expertise:** The scientific experts[2] are established or emerging leaders in the relevant research domains and have the necessary expertise and relevant collaborations to guide the development of and/or exploit the research data infrastructure.

- **Technical expertise:** The technical experts[2] have the required expertise to efficiently design and build the research data infrastructure.

- **Sustainability and maintaining relevance:** A compelling plan for the long-term management of the data is in place to ensure ongoing relevance of the infrastructure. The proposal presents a credible plan addressing the long-term financial sustainability of the research data infrastructure.

- **Benefits to Canadians:** The research activities enabled by the infrastructure have the potential to lead to significant tangible benefits for society, health, the economy and/or the environment. The use of the research data infrastructure will be maximized by adopting best practices in accessibility, interoperability and generalizability.

Multidisciplinary Assessment Committee review

This stage of the merit review involves the assessment of the proposals by a Multidisciplinary Assessment Committee. Following a careful analysis of the results of the Expert Committee review, the MAC will be responsible for selecting proposals that best meet the objectives of the competition and for establishing the amount that should be awarded to each proposal.

---

[2] See section 3 for definition

The MAC will only review proposals that were recommended for funding by the expert committees and will evaluate the proposal in accordance with the following objectives:

- **Scientific excellence:** The proposal relates to a field in which Canada is recognized for having significant research strengths. The research data infrastructure project addresses opportunities or gaps identified by the research community and will enable established and emerging leaders to remain internationally competitive;

- **Impact and ongoing relevance:** A critical mass of Canadian scientific experts is actively involved throughout the project. Use of the research data infrastructure will be maximized through optimal access mechanisms. The research enabled has the potential to lead to significant tangible benefits to society, health, the economy and/or the environment;

- **Feasibility:** The research data infrastructure project's scope is clearly defined and the technical experts have the necessary expertise and experience to design, develop and deliver the project within 36 months. The proposed budget, as well as the long-term financial sustainability of the infrastructure is realistic.

MAC members are chosen for their capacity to assess proposals based on the competition objectives and for their breadth of understanding of the data-intensive research environment and will include expertise in the building and management of research data infrastructure resources.

Decisions

The recommendations of the MAC for the first competition will be submitted to the CFI Board of Directors for consideration at its March 2018 meeting. The CFI Board of Directors makes all funding decisions on CFI awards. The CFI is seeking innovative projects and will endeavor to support a portfolio of five to ten projects that include both larger-scale mature initiatives and smaller-scale emerging initiatives.

Collaboration with provinces and territories and other funding partners

The CFI encourages institutions to work with relevant provincial and territorial funding authorities and other key partners at an early stage in the planning and development of their proposals under this initiative.

To coordinate the review processes and avoid duplication of efforts, the CFI will share a list of the NOIs submitted, along with Expert Committee reports and the names and affiliations of committee members, with relevant provincial and territorial funding authorities identified as funding partners in the proposals. Disclosure of the list and committee reports will only be made in accordance with agreements between the CFI and provincial or territorial authorities, as permissible pursuant to the Privacy Act.

In addition, and where feasible, representatives of the relevant provincial or territorial authorities will be invited to participate as observers at the expert review stage. These representatives will have the opportunity to submit their views on proposals. These views will be considered by the MAC.