Canada Foundation for Innovation Cyberinfrastructure Initiative

Call for Proposals



TABLE OF CONTENTS

1.	CONTEXT	2
	Advanced research computing	2
	Research is increasingly data intensive	2
	Research data infrastructure	2
	The way forward: adopt, adapt and develop	3
	A view of the long term	3
2.	THE CHALLENGE	3
	Challenge 1: to institutions and communities of researchers	3
	Challenge 2: to Compute Canada	4
3.	FUNDS AVAILABLE	5
	Operating and maintenance costs	5
CC	DMPETITION TIMELINES	6
4.	ELIGIBILITY	7
	Institutional eligibility	7
	Research data infrastructure	7
	Eligible infrastructure projects	7
	Ineligible costs under Challenge 1	7
	Partner contributions	7
	Consultative process with Compute Canada	8
	Access to the research data resources	8
5.	REVIEW AND DECISION MAKING	9
	Challenge 1: research data infrastructure project proposals	9
	Expression of interest	9
	Notices of Intent	9
	Review of full proposals	10
	Review Process	10
	Challenge 2: proposals received from Compute Canada	11
	Review process	12

1. CONTEXT

Many of today's research endeavours are driven by massive digitization initiatives, high-throughput devices, sensor platforms and computational modelling and simulation, all of which generate data that are unprecedented in size and complexity. Advanced research computing infrastructure and the accompanying human and capital resources – generically termed cyberinfrastructure – are required to support and catalyze these endeavours. The CFI has long recognized that cyberinfrastructure underpins leading-edge research in all fields of inquiry from the social sciences and humanities to the health and physical sciences through to engineering. Over the past 16 years, the CFI has invested some \$360 million in cyberinfrastructure, for a total investment of about \$900 million when partner funding is taken into account.

Advanced research computing

In 2006, the CFI invested \$78 million in the creation of a pan-Canadian high performance computing platform. This platform was designed to provide internationally competitive, mid-range, advanced computing resources and data storage to researchers who conduct computationally challenging research, irrespective of their location. Subsequently, a consortium of Canadian universities created Compute Canada to manage and coordinate the resources that make up the platform. In doing so, Compute Canada's mandate is to ensure that researchers have access to advanced research computing facilities and the expert services required to remain internationally competitive. Since 2006, the CFI has continued to focus many of its investments on common or shared computational and data storage capabilities that are beyond what can be provided by individual institutions. More recently, the CFI has extended its support to the pan-Canadian computing platform by contributing to its operating and maintenance costs through Compute Canada. This additional support is designed to help the platform fully exploit its capabilities.

Research is increasingly data intensive

Today, research across all fields of inquiry is increasingly data intensive. Data is both 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.

Such vast quantities of data also create significant pressure on Canada's existing advanced research computing infrastructure. Leading-edge, data intensive research will only be successful in generating new knowledge if it is supported by robust research data infrastructure and advanced research computing infrastructure.

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. It will require partnerships within and, in some cases, across research domains in order to foster the development of shared infrastructure that enables community-wide approaches and solutions for a wide range of users. Such partnerships will also reduce unnecessary duplication of efforts. Adopting a collaborative, partnership-based approach will require a collective effort

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.

wherein communities of researchers in particular domains work closely with Compute Canada, with experts that are able to build research data infrastructures (such as data scientists, data analysts and software developers) and with other stakeholders. In many cases, the expertise required to build research data infrastructures already exists within Compute Canada and can be employed effectively. Ultimately, the CFI's goal is to support research communities in their efforts to devise optimal ways to organize and use research data resources and put in place the necessary computational and data storage capabilities.

The way forward: adopt, adapt and develop

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 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, opportunities to grow existing initiatives in Canada should also be considered.

A view of the long term

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.

2. THE CHALLENGE

The overarching objective of the CFI's 2015 Cyberinfrastructure Initiative is to enhance the capacity of Canadian institutions and researchers to conduct leading-edge research in areas of demonstrated strength by supporting the infrastructure needs of computationally- and data-intensive research. The CFI will support these needs by investing in:

- A limited number of research data infrastructure projects that, in collaboration with Compute Canada, enable communities of researchers, along with data scientists, data analysts, software developers and other experts to devise optimal ways of organizing and using research data resources; and.
- Upgrading and modernizing the computational and data storage capacities of the pan-Canadian advanced research computing platform, managed by Compute Canada.

The CFI, therefore, is issuing two interrelated challenges to institutions and researchers.

Challenge 1: to institutions and communities of researchers

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.

The projects are expected to involve multi-institutional consortia of institutions and their researchers, data scientists, data analysts and software developers. These consortia will design and construct research data infrastructure projects that are required by a community of researchers to advance their research programs. The projects must be implemented within a reasonable time frame. Therefore, consortia are encouraged to embrace an "adopt, adapt and develop" approach as well as to link to international initiatives whenever appropriate. This will promote efficiency, interoperability and rapid implementation. Projects that propose to expand and extend existing Canadian data initiatives will also be considered.

As well, the CFI challenges consortia to develop proposals that demonstrate the sustainability and long-term relevance and usefulness of the research data infrastructure.

The CFI requires that the computational and data storage infrastructure needed by the projects be managed by Compute Canada. This will promote the efficient, effective and economical use of the computation and data storage infrastructure managed by Compute Canada. Proponents of the research data infrastructure projects must, therefore, engage Compute Canada to determine if the platform currently has the capabilities required by the project or if new capabilities need to be added to the platform. Funding for computation and data storage required for funded research data infrastructure projects will be provided under Challenge 2.

The CFI is seeking innovative projects and will endeavor to support a portfolio of projects that include both larger-scale mature initiatives and smaller-scale emerging initiatives. The CFI expects to support between five and ten research data infrastructure projects per competition.

Challenge 2: to Compute Canada

Canada's advanced research computing platform is reaching the limits of its capacity. As well, the computational services currently available are not necessarily designed to support tomorrow's leading-edge research. At the same time, Compute Canada is being called upon to provide new services and support to an increasingly diverse user base. Some researchers need secure environments because of their use of confidential information, others need cloud computing services and still others need state-of-the-art high-performance computing resources. In each case, users require the support of experts to assist them in devising effective approaches for their activities as well as user-friendly access regimes. This includes meeting the needs of the research data infrastructure projects as described under Challenge 1.

These needs cannot be addressed by simply replacing existing computational and data storage infrastructure. Enhancing existing capacity in an optimal way requires that priorities be set and choices be made. Over the last year, Compute Canada has consulted with the user community to develop a five-year strategic plan that will guide the activities of the organization. The CFI expects this five-year plan to be the starting point of a pan-Canadian roadmap that will identify the needs of the research community and serve as a basis for the renewal of the advanced computing research platform. This roadmap should also capitalize on the unique strengths of each of the four regional organizations: ACEnet, Calcul Québec, Compute Ontario and WestGrid. In order to fulfil its mandate, Compute Canada will need to adopt a pan-Canadian perspective that ensures that both established and new users of the advanced research computing platform have ready access to the computational capabilities and support services they require irrespective of their institutional affiliation.

3. FUNDS AVAILABLE

The CFI will provide up to \$50 million for this initiative. Following standard practice, the CFI will contribute up to 40 percent of the eligible costs of a project.

The \$50 million will be allocated as follows:

• Challenge 1:

Up to \$20 million in total will be provided for two competitions of approximately equal size for research data infrastructure projects (the competition budget includes the associated Infrastructure Operating Fund). The proposals under Challenge 1 can request up to \$2 million from the CFI Successful proposals will also receive a contribution under the Infrastructure Operating Fund equivalent to 30% of the Challenge 1 award.

Challenge 2:

- Stage 1: Up to \$15 million will be provided for the upgrading and modernization of the computational and data storage capacity of the pan-Canadian advanced research computing platform. As the managers of this platform, Compute Canada will be invited to submit a proposal on behalf of the advanced research computing community; and
- Stage 2: At least \$15 million will be provided for both the computational and data storage needs of the research data infrastructure projects selected under Challenge 1 and the continued upgrading and modernization of the pan-Canadian advanced research computing platform. Compute Canada will be invited to submit a second proposal on behalf of the advanced research computing community.

Operating and maintenance costs

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 host institutions.

The awards for the upgrading and modernization of the advanced research computing platform under Challenge 2 will not be eligible for IOF support because the CFI is already contributing to the O&M costs of the advanced research computing platform through the Major Science Initiatives (MSI) award managed by Compute Canada.

COMPETITION TIMELINES

Key dates	Challenge 1: Research data infrastructure proposals	Challenge 2: Compute Canada
November 28, 2014	1st competition call for proposals released	Call for the Stage 1 proposal to address pressing immediate needs
February 27, 2015	1 st Competition deadline for submission of expressions of interest	
April 17, 2015		Submission of the Stage 1 proposal
May 22, 2015	1 st competition deadline for notices of intent	
June 30, 2015	1 st competition invitation to submit full proposal	CFI DECISION on Stage 1 proposal
October 16, 2015	1st competition deadline for full proposals; 2nd competition call for proposals released	
March 2016	CFI DECISION on 1 st competition proposals	
April 2016	2 nd 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	2 nd competition invitation to submit full proposals	
June 2016		CFI DECISION on Stage 2 proposal
October 2016	2 nd competition deadline for full proposals	
March 2017	CFI DECISION on 2 nd competition proposals	

4. ELIGIBILITY

Institutional eligibility

Canadian universities, colleges, research hospitals and non-profit research institutions that have been recognized as eligible by the CFI can apply. For proposals submitted under Challenge 2, the administrative institution must be a member of Compute Canada as well as a CFI-eligible institution.

Research data infrastructure

A research data infrastructure refers to a research resource and associated software and analytical tools. The research data resource consists of data and metadata that is structured, organized and standardized through open, documented, common and interoperable schemas, protocols and formats. Under the Cyberinfrastructure Initiative's Challenge 1, it does not include computational hardware or data storage, which are addressed under Challenge 2.

Eligible infrastructure projects

The existing CFI definition of eligible infrastructure will apply. The CFI will contribute to these costs up to the point where the research data infrastructure is ready to be used by researchers. For added clarity, under Challenge 1, 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;
- 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; and
- the renovation of existing space, for the personnel required to develop and build the research resource.

Ineligible costs under Challenge 1

The costs for computational hardware and data storage are not eligible under Challenge 1. These shared computational and storage resources will be part of the pan-Canadian advanced research computing platform, and therefore these resources will be requested under Challenge 2, as part of an integrated proposal submitted by Compute Canada. The Challenge 2 proposal to be submitted in April 2016 will specifically address the needs and requirements of successful proposals under Challenge 1.

The costs for primary data collection are also ineligible.

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 for datasets that are not otherwise sold to third parties, 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, 2015. Expenditures are considered incurred when goods are received, services have been rendered or work has been performed.

Consultative process with Compute Canada

The CFI expects that the projects funded under Challenge 1 will likely require similar capabilities and share common needs, and therefore economy, efficiency and effectiveness will be better served by making Compute Canada responsible for the development of the optimal technical solution and configuration to specifically address the needs and requirements of successful proposals under this challenge. As a result, computational hardware and data storage equipment will not be eligible costs under Challenge 1.

Compute Canada's resources and efforts would be best utilized in discussions once the initiative is sufficiently mature and progressed beyond the EOI. As such, there is no requirement or expectation for applicants to hold formal discussions with Compute Canada in the early stages of development of their project. However, the CFI expects, that discussions will begin between applicants and Compute Canada at the NOI development stage, although there will be no formal requirements regarding these discussions in the NOI.

Finally, the CFI will require that all those invited to submit full proposals will hold formal discussions with Compute Canada to determine if the pan-Canadian platform 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 — existing or new. These requirements will form the basis for the Challenge 2 proposal to be submitted by Compute Canada in April 2016.

Access to the research data resources

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, whether or not he or she is affiliated with an institution that is a member of the consortium.

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.

5. REVIEW AND DECISION MAKING

Challenge 1: research data infrastructure project proposals

Expression of interest

The CFI invites the applicant institutions who will act as administrative host to submit expressions of interest for research data infrastructure projects, indicating whether the project is being considered for the first or second competition. The expressions of interest should include the following: project title, composition of the consortium (institutions and principal members of the consortium), a one paragraph summary (200 words) describing the proposed research data infrastructure project, the estimated total project cost and the amount of funding that will be requested from the CFI. Expressions of interest should be submitted **no later than February 27, 2015** at cyber@innovation.ca.

The expressions of interest will be used by CFI staff to gauge the level of interest in this competition and to plan the competition requirements. In order to help others to identify opportunities and encourage collaboration between consortia in closely-related or identical domains, the full list of expressions of interest will be published on the CFI website.

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 full 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). Using assessment criteria that focus on the scientific merit of the project and its maturity, this merit review will identify those proposals that best fit the objectives of the Cyberinfrastructure Initiative.

Under this Call for proposals, the CFI anticipates receiving approximately 50 NOIs. The full list of NOIs will be posted on the CFI website. The CFI will invite full 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 of the NOIs in which it participates.

The MAC will assess NOIs on the basis of the following criteria:

- Canada has significant research strengths and globally competitive expertise in the area of the proposal;
- The research data infrastructure project is essential in maintaining Canada at the leading-edge internationally;
- The consortium is composed of a critical mass of leading Canadian researchers in the domain who are fully engaged in the project and who are capable of exploiting the full potential of the research data infrastructure;
- The consortium has the required expertise to design and construct the research data infrastructure and develop the analytical tools, methods and standards to achieve the desired outcomes; and,
- The scope of the project has been clearly defined and allows for commissioning within two to three years.

Review of full proposals

The full proposals should clearly articulate the merits and excellence of the proposed project and provide sufficient information to enable reviewers to evaluate the proposal in accordance with the following criteria:

- **Research or technology development:** The proposed research activities enabled by the research data infrastructure are timely, innovative and at the leading edge internationally.
- Consortium: The consortium has the required expertise to design and build the research data infrastructure resource and to develop the analytical tools, data management methods and standards required to enable the proposed research activities. As well, the consortium members who are the principal users of the research data infrastructure are established or emerging leaders in the relevant research domains. The team has the necessary expertise, ability and relevant collaborations and partnerships in place to exploit the full potential of the research data infrastructure.
- **Infrastructure:** The research data infrastructure is necessary and appropriate to conduct the proposed research activities. The scope and requirements of the project, including development of tools, methods and standards are well defined and able to be commissioned within two to three years.
- Sustainability and maintaining relevance: The proposal presents a credible and convincing
 plan that addresses the long-term financial sustainability of the research data infrastructure. A
 compelling plan for the long term management of the data is in place to ensure ongoing
 relevance and usefulness of the 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.

Review Process

The assessment of each full proposal will be conducted in two stages:

Expert review

In the first stage of the merit review, expert committees will be convened to assess the strengths and weaknesses of the proposals in relation to the five assessment criteria for this competition. These committees will be tasked with recommending to the Multidisciplinary Assessment Committee (MAC) those proposals that meet the standard of research 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. Large and/or complex proposals may involve face-to-face meetings between the expert committee members and project proponents, including senior representatives of the participating institutions.

Additional written external reviews may be solicited and used by the expert committees if supplementary expertise is required.

Multidisciplinary Assessment Committee (MAC)

The second stage of the merit review involves the assessment of the full proposals by a Multidisciplinary Assessment Committee (MAC). Following a careful analysis of the results of the expert committee review, the MAC will be responsible for:

- Selecting proposals that best meet the standards of excellence for the competition; and,
- Establishing the amount that should be awarded to each proposal.

The MAC will only review proposals that were recommended for funding by the expert committees. The MAC will use the same assessment criteria as were used at the NOI stage.

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 2016 meeting. The CFI Board of Directors makes all funding decisions on CFI awards.

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.

Challenge 2: proposals received from Compute Canada

Stage 1: The renewal of the pan-Canadian advanced research computing platform will be conducted in two stages. For Stage 1, the CFI invites Compute Canada, on behalf of its member institutions, to propose three distinct options for the capabilities and services that will enable leading-edge research and address the most pressing immediate needs. This proposal will focus on the upgrading and modernizing of the computational and data storage capabilities managed by Compute Canada.

Stage 2: Once the results of the competition for research data infrastructure projects are known, the CFI will invite Compute Canada to develop and submit a second proposal containing three distinct options that aims to optimize the computational and data storage requirements of the funded

research data infrastructure projects under Challenge 1 and to continue the upgrading and modernizing of the pan-Canadian advanced research computing platform.

The CFI expects that both Stage 1 and 2 proposals will be developed in close collaboration with the advanced research computing community, as part of a pan-Canadian strategy that will clearly identify the priorities and choices made to optimize its offerings. The proposals should also capitalize on the unique strengths of each of the four regional organizations: ACEnet, Calcul Québec, Compute Ontario and WestGrid.

Proposals for both stages should clearly articulate the merits and excellence of the research that will be enabled. In addition, they should provide sufficient information to enable the expert committee to evaluate the proposals in accordance with the following criteria:

- Research or technology development: The proposed research or technology development
 activities enabled and supported by the research infrastructure are timely, innovative and at the
 leading-edge internationally.
- **User access:** The proposal presents a plan that will provide effective and efficient support to users with an excellence-based access regime.
- Infrastructure: The proposed infrastructure is state-of-the-art, technologically appropriate, well-aligned with the strategic plan of Compute Canada and addresses the current and anticipated needs of the Canadian advanced research computing community. The proposed infrastructure will lead to strategic enhancements of the platform through rationalization and optimization of the computational capabilities and services offered by Compute Canada.
- **Sustainability:** The proposal presents a compelling plan for the management, operation and maintenance of the proposed infrastructure with tangible and appropriate commitments over its useful life.
- **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.

Review process

Expert committee

Because the platform renewal will have a direct impact on the operations and maintenance (O&M) budget of Compute Canada, the review of the Stage 1 proposal will be conducted in conjunction with the mid-term review of the Major Science Initiatives award managed by Compute Canada. This approach is consistent with a life-cycle approach and will enable the experts to simultaneously review the proposed computational and data storage needs and the O&M needs of the advanced research computing platform. This integrated review will enable a thorough evaluation of the appropriateness and viability of the proposed enhancements.

The expert committee will evaluate the strengths and weaknesses of the proposal, and make a recommendation on funding. If the expert committee recommends funding, it will be asked to establish an appropriate amount.

The CFI anticipates that Compute Canada will submit the Stage 1 proposal in April 2015 and the Stage 2 proposal in April 2016. Both proposals will be reviewed by the same committee. In each case, the review will involve a face-to-face meeting between the expert committee members and project proponents, including principal users and senior representatives of the participating institutions.

Decisions

The expert committee will submit its recommendation for the Stage 1 proposal to the CFI Board of Directors for consideration at its June 2015 meeting. For the Stage 2 proposal, the expert committee will submit its recommendation to the CFI Board of Directors for consideration at its June 2016 meeting. The CFI Board of Directors makes all funding decisions on CFI awards.