2023-08-25

National Research Software Strategy 2023

National Research Software Strategy Working Group

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

Table of Contents

Executive Summary and Purpose

In February 2016, the Minister of Science, the Honourable Kirsty Duncan, met with members of the Leadership Council for Digital Research Infrastructure (LCDRI) and asked them to affirm their intention to work collaboratively to ensure that Canada has a strong DRI ecosystem in place to support the critical research undertaken by Canada's academic research community. In November 2016, Innovation, Science, and Economic Development Canada (ISED) provided funding to the LCDRI to establish a short-term secretariat to develop position papers on data management (DM); advanced research computing (ARC); and recommendations for potential approaches for coordination of national DRI. In its submission to ISED, the LCDRI listed the five core components of DRI as "*network, advanced research computing, data management, storage, and advanced research software*",[1] recognizing Research Software (RS) as a key pillar. At the same time the LCDRI recognized that research software was not considered as part of its original remit: "*... while the development and stewardship of advanced research software (ARS) is represented in all options, consideration of how best to coordinate and deliver this component was not part of LCDRI's current remit*". As a result, although the LCDRI produced position papers for Advanced Research Computing (ARC) and Research Data Management (RDM) no such position paper was produced for RS. [2]

In 2018 the Government of Canada invested $572.5 million over 5 years (Budget 2018) to enhance Canada's DRI ecosystem. [3] This led to a 2019 proposal to ISED for the creation of a new organization to coordinate funding and strategic directions for national activities related to ARC, RDM and RS. [4] This proposal was approved and the Digital Research Alliance of Canada (or the Alliance, formerly New Digital Research Infrastructure Organization, or NDRIO) was launched in March 2020. Between 2020 and 2021, the Alliance and its stakeholder community developed current state assessments for ARC (Current State of Advanced Research Computing in Canada), RDM (The Current State of Research Data Management in Canada), and RS (Research Software Current State Assessment).[5] The *Research Software Current State Assessment* was the first position paper on RS in Canada, and was written with the goal of

---

[1] LCDRI Coordination Position Paper (unpublished), Submitted to ISED 2017

[2] LCDRI Advanced Research Computing (ARC) Position Paper, Unpublished Report, 2017-08-31; LCDRI Data Management Position Paper, Unpublished Report, 2017-08-31

[3] "Canada: Budget 2018", archived, https://www.budget.gc.ca/2018/home-accueil-en.html

[4] "Digital Infrastructure Contribution Program Program Guide", https://ised-isde.canada.ca/site/digital-research-infrastructure/sites/default/files/attachments/DRIContributionProgram_ProgramGuide.pdf

[5] Sahrakorpi et al., "Current State of Advanced Research Computing in Canada", Zenodo, May 31, 2021, https://doi.org/10.5281/zenodo.6547111; Khair et al., "The Current State of Research Data Management in Canada", Zenodo, November 30, 2020, https://doi.org/10.5281/zenodo.6564658; Zhang et al., "Research Software Current State Assessment", Zenodo, June 21, 2023, https://doi.org/10.5281/zenodo.8066970

closing the gap between RS and RDM and ARC. From 2021 to 2022, the Alliance transformed from start-up to scale-up, accompanied by a series of transition activities of the national mandates for RDM, ARC and RS to the Alliance. For RS in particular, both CANARIE's RS program and some of CFI's initiatives have been transitioned to the Alliance. The Alliance has assumed the full leadership for the RS mandate in Canada, as a national funder, service provider and coordinator.

The Alliance's Strategic Plan 2022-2025 highlights the importance of RS: [6]

> *The Alliance will: Support researchers in adopting best practices of research software development and management throughout the research lifecycle through targeted funding opportunities, training, service catalogues, policies and community development. The Alliance values and recognizes research software as a first-class research output.*

Despite these efforts, the RS initiative "Towards a Sustainable Canadian Research Software Community" put forward to ISED as part of the Alliance's Multi-Year Funding Proposal (MYFP) for 2023-25 was not approved for funding, leaving a significant gap in the development of RS capability, RS community, and RS coordination in Canada.

The Alliance RS Strategy for 2025-2030 (the RS Strategy) builds on past RS foundations, leverages the strengths of past RS programs, while at the same time addressing the weaknesses of having a historically disjoint and uncoordinated national strategy for RS. The RS Strategy presents:

- an overview of the current state of the RS landscape nationally and internationally;
- a set of challenges and opportunities in RS service provision for the Alliance as a national funder, service provider and coordinator;
- a vision and goals to support the complete research software (RS) lifecycle (developing, deploying, and sustaining) while also developing career paths for the research software experts (RSEs) who create and support it; [7]
- strategic recommendations that will increase the value of an integrated DRI ecosystem by focusing on high-value aspects of RS and in particular where RS overlaps with ARC and RDM.

[6] Digital Research Alliance of Canada, "Alliance Strategic Plan 2022-2025", https://stratplan.alliancecan.ca/wp-content/uploads/2022/03/Alliance_Strategic_Plan_2022_2025.pdf

[7] Research Software Expert is used in Canada as Engineer is a protected term: the international community uses the phrase Research Software Engineer.

The RS Strategy is authored by the RS Strategy Working Group consisting of the following members: Mark Leggott (Chair), Girma Bitsuamlak, Susan Brown, Liseanne Cadieux, Brian Corrie, Carolyn Côté-Lussier, Fares Dhane, Félix-Antoine Fortin, Kimberley Hartley, Tamanna Moharana, Pierre-Olivier Quirion, Y.G. Rancourt, Seppo Sahrakorpi, Abdel Yousif, and Qian Zhang.

RS Vision for 2025-2030

To build a world-class, equitable, and sustainable research software ecosystem in Canada through advancing capability, community, and coordination, delivering research software that spans the DRI ecosystem and enables advances in research and innovation.

RS Strategic Goals for 2025-2030

1. **Build Capability** by ensuring that Canada has the capability/capacity to **develop** and **sustain** world-class **RS Tools** and **RS Platform**s.
2. **Build Community** by developing an active and skilled Canadian community (**HQP**) around the **use** and **development** of RS.
3. **Coordinate** with stakeholders to provide **governance** of the RS **ecosystem** and **establish** Canadian RS **policy.**

Summary of RS Recommendations for 2025–2030

The RS Strategy was developed by the RS Strategy Working Group (RSSWG) from January to July 2023. This document draws on a wide range of sources, including data from national and international organizations, international funder presentations, and the extensive background material collated by the Alliance on user needs, stakeholder feedback, user surveys, and the current state assessments.[8]

The 2025-2030 recommendations are split into a set of categories (pictured) within each of the high-level goals of advancing RS **Capability, Community**, and **Coordination** in Canada (see

[8] https://alliancecan.ca/en/initiatives/canadian-digital-research-infrastructure-needs-assessment; https://alliancecan.ca/en/initiatives/position-paper-submissions; https://alliancecan.ca/en/latest/news/what-are-your-thoughts-research-software-complete-our-survey-june-2; https://alliancecan.ca/en/initiatives/current-state-reports

Figure 1). We describe these categories at a high-level here, with a detailed description of the recommendations in each of these categories given in the Recommendations section.

1. **Capability**: **RS Capability in Canada is substantial and yet fragile**. Canadian RS capabilities are substantial, despite the ad-hoc and uncoordinated RS funding approach and RS strategy in Canada to date. As a result, these important platforms are fragile and at risk. The Alliance should develop RS **Funding** approaches that ensure that world class RS developed by Canadian teams is sustainable while at the same time new and innovative RS can be developed. In addition, the Alliance needs to develop a national RS middleware **Infrastructure** capability (authentication, data discovery, job management, security, etc.) to support Canadian researchers that are developing such RS. Finally, RS capabilities and support at the intersection of RS and **Research Data** and **Cybersecurity** need to be enhanced to ensure that RS development teams are utilizing and following RDM and Cybersecurity protocols and best practices when developing RS.
2. **Community**: **A strong RS community is critical to the DRI ecosystem**. RS plays a key integrative role across the RDM and ARC pillars. The Alliance should work with national stakeholders towards the development of an RS **Training** program that is integrated with ARC and RDM, as well as providing in-depth **Support** to researchers in finding and developing RS solutions to their research problems across the DRI ecosystem. In addition, the Alliance should work with stakeholders to create community driven **Events** that bring together the RS, ARC, and RDM communities, breaking down barriers and increasing the interaction across the entire DRI ecosystem. Community building in all of its forms should be looked at to ensure equity, diversity, inclusion, and accessibility (EDIA).
3. **Coordination**: **RS Policy and Governance is immature in Canada, and requires coordination with all members of the stakeholder community to achieve a sustainable approach**. It is critical that the Alliance collaborate with stakeholders in the national and international DRI Community to develop and adhere to **Policy** and **Governance models** that will help to raise RS to a "first-class" output of research in Canada. The Alliance should work with national (e.g., institutions, regional organizations, funders, research community) and international stakeholders (e.g., Research Software Alliance (ReSA), international funders, international RS societies) to develop Canada's RS HQP community, ensure that stakeholder policies consider and support RS, and facilitate the development of policy and governance that will advance RS as a critical pillar in the DRI ecosystem nationally and internationally.

Figure 1. Recommendation categories, grouped across the goals of building RS Capability, Community, and Coordination.

What is Research Software?

The FAIR4RS principles consider RS as a class of digital object, extending the FAIR principles for Research Data to include specific aspects of RS that are critical to making Research Software (rather than Research Data) Findable, Accessible, Interoperable, and Reusable (FAIR).[9] The FAIR4RS principles defines RS as "*source code files, algorithms, scripts, computational workflows and executables that were created during the research process or for a research purpose. Software components (e.g., operating systems, libraries, dependencies, packages, scripts, etc.) that are used for research but were not created during or with a clear research intent should be considered software in research and not research software.This differentiation may vary between disciplines.*"

Although this definition of RS is simple, it can be broadly interpreted. The DRI ecosystem in which RS is used is complex and multidimensional - RS is critical to the vast majority of modern research, RS varies dramatically across disciplines, RS relies on a disparate and diverse development community for its success, RS is fundamental to how research data is accessed and managed (RDM), and the development of RS is an essential component to advancing research on ARC platforms. The FAIR4RS definition above is cited from the aptly named paper "Defining Research Software: a controversial discussion",[10] where the definition also includes the phrase "*The minimal requirement for achieving computational reproducibility is that all the computational components (Research Software, software used in research, documentation and hardware) used during the research are identified, described, and made accessible to the extent that is possible.*" The line between RS and "software used in research" is difficult to define and often changes based on the maturity of the software in question. For example, the programming language R was originally developed as an RS project but today is more aptly considered "software used in research" as it is a mature, sustainable, open source package that many researchers rely on for data analysis.[11] RS as discussed in the context of the RS Strategy is inclusive of both of these categories of software, recognizing that both RS and "software used in research" as defined above are critical to the DRI ecosystem and require a strategic approach to their support and development.

---

[9] Barker et al., "Introducing the FAIR Principles for research software", *Sci Data* 9, 622 (2022), https://doi.org/10.1038/s41597-022-01710-x; Wilkinson et al., "The FAIR Guiding Principles for scientific data management and stewardship", Sci Data 3, 160018 (2016), https://doi.org/10.1038/sdata.2016.18

[10] Gruenpeter et al., "Defining Research Software: A Controversial Discussion". Zenodo, September 13, 2021. https://doi.org/10.5281/zenodo.5504016.

[11] Ihaka et al., "R: A Language for Data Analysis and Graphics", *Journal of Computational and Graphical Statistics* (1966) 5 (3): 299–314, https://doi.org/10.1080/10618600.1996.10474713; The R Project for Statistical Computing, https://www.r-project.org/

As discussed above, RS plays multiple roles in research, with the Alliance RS Current State Assessment classifying RS into two categories: [12]

- RS Tools: RS as a broad term may include software programs, languages, libraries, scripts, computational code, models, electronic lab notebooks, repository software, workflow management and other tools developed to support research or facilitate research processes, which themselves can be implemented in RS. RS tools can be packaged and shared as source code, executables, containers, services, etc.
- RS Platforms: As research infrastructure, RS often refers to RS platforms or services (variously referred to as virtual science labs, virtual research environments or VREs, or Science Gateways) that instantiate a collection of RS source code, including complex software frameworks, discipline-specific tools, services, and glue code.[13] These are deployed to support both the research workflows and the communities of practice engaged in collaborative research (e.g., Syzygy, CWRC).[14] Typically, RS platforms span the ARC and RDM DRI pillars, with capabilities that include one or more of data acquisition and management, processing and visualization, storage and preservation, sharing and discovery (e.g., Federated Research Data Repository (FRDR)[15]); platforms may provide the full spectrum or a subset of components. RS platforms may be discipline-specific and may support and enhance scientific collaboration and scholarly communication by facilitating citizen science engagement as well.

See the Alliance's RS Current State paper for a more detailed description of RS. [16]

---

[12] Zhang et al., "Research Software Current State Assessment". Zenodo, June 21, 2023. https://doi.org/10.5281/zenodo.8066970.

[13] Barker et al., "The global impact of science gateways, virtual research environments and virtual laboratories", *Future Generation Computer Systems* 95 (2019): 240-248, https://doi.org/10.1016/j.future.2018.12.026.

[14] https://syzygy.ca/; https://cwrc.ca/.

[15] https://www.frdr-dfdr.ca/repo/

[16] Zhang et al., "Research Software Current State Assessment". Zenodo, June 21, 2023. https://doi.org/10.5281/zenodo.8066970.

RS Value Proposition

Research software has been recognized internationally as a key pillar in the support of Open Scholarship and the DRI Ecosystem. [17] The Canadian Roadmap for Open Science defines Open Science as "*The practice of making scientific inputs, outputs and processes freely available to all with minimal restrictions*". [18] The 2021 UNESCO definition of Open Scientific Knowledge is broad and inclusive, stating that "*Open scientific knowledge refers to open access to scientific publications, research data, metadata, open educational resources, software, and source code and hardware that are available in the public domain…*"[19]

Research software is ubiquitous in research and foundational research software (e.g., Matplotlib, Numpy, Pandas) is critical to accelerating science.[20] Managing research software is crucial for academic institutions and research organizations to enhance their productivity, collaboration, and overfall scientific impact. A recent study of researchers from Russell Group Universities (24 research intensive universities) in the UK showed that 92% of researchers used software in their research, with 70% stating that it would be impractical to do their research without software.[21] 56% of those surveyed stated that they developed their own software. Similar results were obtained in a recent study in the US, with 95% of respondents stating they used software and 63% stated they could not do their research without software.[22] The Alliance's Needs Assessment reflected similar results with a significant disparity between research domains: "*In SE [Science and Engineering], 80% of respondents write code/software,*

---

[17] In this document, we use the term Open Scholarship rather than Open Science to be more inclusive of the Social Sciences and Humanities, except where the term "Open Science" is used directly in a title, a quote, or a paraphrase. From Tenant et al. "The term 'open science' implicitly seems only to regard 'scientific' disciplines, whereas open scholarship can be considered to include research from the Arts and Humanities" - Tenant et al., A tale of two 'opens': intersections between Free and Open Source Software and Open Scholarship, https://doi.org/10.31235/osf.io/2kxq8

[18] https://science.gc.ca/site/science/en/office-chief-science-advisor/open-science

[19] https://en.unesco.org/science-sustainable-future/open-science/recommendation, Section II.7.d

[20] Hunter et al., "Matplotlib: A 2D Graphics Environment", *Computing in science & engineering* 9, no. 03 (2007): 90-95, https://doi.org/10.1109/MCSE.2007.55; Harris et al., "Array programming with NumPy", *Nature* 585, no. 7825 (2020): 357-362, https://doi.org/10.1038/s41586-020-2649-2; The pandas development team. "Pandas-dev/pandas: Pandas", Zenodo, October 26, 2023. https://doi.org/10.5281/zenodo.10045529; Knowles et al., "We need to talk about the lack of investment in digital research infrastructure", *Nature Computational Science* 1, no. 3 (2021): 169-171, https://doi.org/10.1038/s43588-021-00048-5

[21] https://www.software.ac.uk/blog/2014-12-04-its-impossible-conduct-research-without-software-say-7-out-10-uk-researchers, https://zenodo.org/record/1183562.

[22] Nangia et al., "Track 1 Paper: Surveying the U.S. National Postdoctoral Association Regarding Software Use and Training in Research", In *Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE 5.1)*. 2017, https://doi.org/10.6084/m9.figshare.5328442.v3

test code/software (67%) [...], yet these activities are much less prevalent from respondents in both SSH [Social Sciences and Humanities] (36%; 24%[..] respectively) and HR [Health Research] (55%; 39% [...] respectively)."[23] Similarly, the Alliance RS Survey showed that 91% of respondents consider RS critical to their research, with 50% considering software a primary output of their research.[24] Despite its ubiquitous nature, research software is typically treated as a less recognized or "*second-class*" research output,[25] making it hard to cite, difficult to measure its impact, and making research reproducibility more challenging.

The international research community has developed thousands of research software applications, some of which have flourished and remain essential components of the research workflows of many researchers, and some of which are no longer accessible.[26] The Alliance RS Survey indicates that 94% of researchers believe software sustainability is important, with software sustainability viewed as important in facilitating data reusability, facilitating research reproducibility, supporting open scholarship principles, and building community. Achieving a balanced approach to funder support is key to ensuring we are addressing both sides of a success equation for research software: the stimulation of new innovation that helps catapult research to the next stage, and the ongoing sustainability of tools that researchers cannot do without.

The challenge in developing and sustaining a strong research software community as part of the DRI ecosystem is well known, with international organizations such as the Research Software Alliance (ReSA) bringing together groups like the Funders Forum to try and address this issue.[27] The Funders Forum is a forum of both organizations committed to supporting research software, and those who develop that software. In a recent example, the 1st International RS Funders Workshop in November 2022 resulted in the Amsterdam Declaration, which is intended to help drive the conversation around sustainable and appropriate funding for research software, starting with those who fund it.[28] The Recommendations in the draft Declaration are a key input into this Strategy. Similarly, the Global Biodata Coalition is "*... a forum for research funders to better coordinate and share approaches for the efficient management and growth of biodata resources worldwide*" with the goal to "*... stabilize and ensure sustainable financial*

---

[23] https://alliancecan.ca/en/document/173

[24] Corrie et al., "Alliance Researcher Council Research Software Survey Report", Zenodo, 2023, https://doi.org/10.5281/zenodo.8280449.

[25] Istrate et al., "A large dataset of software mentions in the biomedical literature", *arXiv preprint arXiv:2209.00693* (2022), https://doi.org/10.48550/arXiv.2209.00693.

[26] Knowles et al., "We need to talk about the lack of investment in digital research infrastructure", *Nature Computational Science* 1, no. 3 (2021): 169-171, https://doi.org/10.1038/s43588-021-00048-5

[27] https://www.researchsoft.org/funders-forum/

[28] Research Software Alliance. "Amsterdam Declaration on Funding Research Software Sustainability", Zenodo, August 1, 2023. https://doi.org/10.5281/zenodo.8325436 (see also https://www.future-of-research-software.org/; https://adore.software/).

support for the global biodata infrastructure" and is working towards identifying "*... a set of Global Core Biodata Resources that are crucial for sustaining the broader biodata infrastructure*".[29] The challenge of developing and sustaining a healthy RS community is a global problem.

When considering how best to support RS, ReSA suggests there are 3 high-level areas of support: the infrastructure; the people who use and create it; the policy that drives and supports it. While all 3 are critical to the RS ecosystem, without strong support for HQP, the software itself will fade and fail to deliver the functionality that researchers need. The 2020 UK R&D Roadmap states:[30]

> *There are a number of issues and challenges we face. Careers in research and development are not as attractive as they should be due to lower salaries and an overdependence on competing for short-term funding. There is an unclear career pathway for many technicians, graduates, early stage researchers and those re-entering research after a break. We do not do enough to support people to move between academia, industry, the public sector and charities, or to move creatively between research and development roles. There are barriers and costs for international researchers and entrepreneurs wanting to come and work here. Creating the right culture is also key.*

As national organizations such as the Australian Research Data Commons (ARDC) have come to realize, having a successful Research Software ecosystem requires having a national agenda that recognizes Research Software as "*... a first-class output of research*",[31] and this approach is strongly aligned with the Alliance's Strategic Plan for RS:[32]

> *The Alliance will: Support researchers in adopting best practices of research software development and management throughout the research lifecycle through targeted funding opportunities, training, service catalogues, policies and community development. The Alliance values and recognizes research software as a first-class research output.*

---

[29] https://globalbiodata.org/

[30] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/896799/UK_Research_and_Development_Roadmap.pdf

[31] Australian Research Data Commons, "A National Agenda for Research Software", Zenodo, March 28, 2022, https://zenodo.org/record/6378082

[32] Digital Research Alliance of Canada, "Alliance Strategic Plan 2022-2025", https://stratplan.alliancecan.ca/wp-content/uploads/2022/03/Alliance_Strategic_Plan_2022_2025.pdf.

Research Software in Context

Approaches to National and International RS Funding

At a high level, there are two categories of funding that are used to develop RS. ReSA defines these two RS funding categories and their key characteristics as follows:[33]

1. **Category A**: **Within a research project as a by-product of doing the research**: Most funding for research software fits into this category, although it is difficult to quantify this directly. Research has shown that from 1995-2016 18,952 awards totalling $9.6 billion (approximately 20% of the NSF research budget) mention software in their abstracts.[34] Other studies have shown that over 90% of researchers use RS, approximately 70% of researchers could not do their research without RS, and 56% of research projects generate new code[35]. Unfortunately, because this software is often produced as a "by-product" of general research, it is often not directly mentioned as an output in funding applications nor reported as a direct research output.
2. **Category B**: **Through intentional development of a software product for general use in research by one or more projects:** Direct funding for research software as the key output of a funding call is provided by some funders. These can be difficult to identify, as they are often embedded in calls for "Cyberinfrastructure", "Open Science", or special, targeted, domain specific calls (e.g., COVID19 data repository platforms).

Research funding also struggles to achieve a balance between the **stimulation of new innovation** that helps catapult research to the next stage and the **ongoing sustainability of tools that researchers cannot do without**.[36] As such, funding calls are often split across these two dimensions. Last, but not least, funding programs sometimes focus on the development and

---

[33] Barker et al., "Overview of Research Software Funding Landscape". Zenodo, February 24, 2022, https://www.doi.org/10.5281/zenodo.6102487.

[34] Katz, "Towards Sustainable Research Software". December 1, 2021, https://doi.org/10.5281/zenodo.5748175

[35] https://www.software.ac.uk/blog/2014-12-04-its-impossible-conduct-research-without-software-say-7-out-10-uk-researchers; Hettrick, "Softwaresaved/software_in_research_survey_2014: Software in Research Survey", Zenodo, February 23, 2018, https://zenodo.org/record/1183562; Nangia et al., "Track 1 Paper: Surveying the U.S. National Postdoctoral Association Regarding Software Use and Training in Research", In *Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE 5.1)*. 2017, https://doi.org/10.6084/m9.figshare.5328442.v3.

[36] Knowles et al., "We need to talk about the lack of investment in digital research infrastructure", *Nature Computational Science* 1, no. 3 (2021): 169-171, https://doi.org/10.1038/s43588-021-00048-5.

sustainability of the RS community (rather than the RS itself), including the development of the critical RS HQP that are central to that community.

An extensive list of international funding programs that were investigated for this report and summarized below are tabulated in Appendix E.

Building Capability

It is extremely challenging to discern quantitatively how much RS is funded through Category A funding programs as described above. Such programs are science driven, and we know that much of that science involves the creation of new software. Because RS is not the direct purpose of the funding, RS creation is rarely presented as an outcome. With initiatives such as the FAIR4RS principles and simple mechanisms to publish software versions in GitHub repositories and acquire DOIs for those software versions, we can expect this to change over time, but to date, funding of this type is extremely difficult to quantify. [37] Many efforts have been made to quantify this, but it remains a challenge. [38] According to NSERC budget and expenditure reporting, NSERC funded approximately $5B in research projects from 2016 to 2021.[39] If we use the 20% factor from the analysis of NSF funding mentioned previously, the $5 B NSERC invested in research between 2016 and 2021 would suggest that a figure of over $1 B in research investments include the development of software. Although this number is substantial, just as it is internationally, it is extremely difficult to quantify in Canada. We therefore focus on direct (Category B above) RS programs to explore the international RS funding landscape. We further split the type of funding across RS Innovation and RS sustainability, with a specific section on Thematic RS funding.

RS Innovation

Most countries have some sort of focussed RS Innovation funding programs, sometimes individual calls and other times large scale RS initiatives. For example, the NIH *Big Data to*

---

[37] Barker et al., "Introducing the FAIR Principles for research software", *Sci Data* 9, 622 (2022), https://doi.org/10.1038/s41597-022-01710-x; https://docs.github.com/en/repositories/archiving-a-github-repository/referencing-and-citing-content

[38] Katz, "Towards Sustainable Research Software". December 1, 2021, https://doi.org/10.5281/zenodo.5748175; Istrate et al., "A large dataset of software mentions in the biomedical literature", *arXiv preprint arXiv:2209.00693* (2022), https://doi.org/10.48550/arXiv.2209.00693.

[39] NSERC: 2022-23 Departmental Plan, ; NSERC: Departmental Results Report 2018-19, https://www.nserc-crsng.gc.ca/NSERC-CRSNG/Reports-Rapports/drr/2018-2019/index_eng.asp#s5.1; NSERC: 2021–22 Departmental Results Report, https://www.nserc-crsng.gc.ca/NSERC-CRSNG/Reports-Rapports/DRR/2021-2022/#s5.1

Knowledge (BD2K) Common Fund Program supported "... *the research and development of innovative and transformative approaches and tools to maximize and accelerate the utility of big data and data science in biomedical research*." [40] This program spanned 2013-2018 and funded 100s of projects and resulted in over 200 RS tools being developed. A snapshot of relatively current national funding programs for RS Innovation is given below, based on the ReSA *Overview of research software funding landscape:*[41]

1. Australian Research Data Commons (ARDC): Platforms co-investment program;[42]
2. Dutch Research Council (NWO): Open science fund;[43]
3. European Commission: Horizon Europe Research infrastructures;[44]
4. Netherlands eScience Center (NLeSC): Open eScience;[45]
5. Nordic e-Infrastructure Collaboration (NeIC): eInfrastructure collaboration;[46]
6. SAGE Publishing: Concept grants;[47]
7. UK Research and Innovation (UKRI): Transformative research technologies;[48]
8. US NSF: Cyberinfrastructure for Sustained Scientific Innovation (CSSI);[49]
9. US National Institutes of Health (NIH): Software tools for open science;[50]
10. Wellcome Trust: Data for science and health;[51]

In Canada, there have been two primary programs that directly support RS. CANARIE's RS program funded RS Innovation from 2007 to 2022, well before Research Software Engineering and its challenges were being recognized internationally. Many of the platforms developed through this successful innovation program are listed on the Alliance's RS Platform directory

---

[40] https://commonfund.nih.gov/bd2k

[41] Barker et al., "Overview of Research Software Funding Landscape". Zenodo, February 24, 2022, https://www.doi.org/10.5281/zenodo.6102487.

[42] https://ardc.edu.au/collaborations/strategic-activities/platforms/

[43] https://www.nwo.nl/en/calls/open-science-os-fund-2020/2021

[44] https://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe/research-infrastructures_en

[45] https://www.esciencecenter.nl/news/the-netherlands-escience-center-presents-new-calls-for-proposals-2021/

[46] https://neic.no/news/2021/01/14/open-call/

[47] https://group.sagepub.com/blog/sage-concept-grants-2022-applications-open

[48] https://www.ukri.org/opportunity/trdf-transformative-research-technologies/

[49] https://beta.nsf.gov/funding/opportunities/cyberinfrastructure-sustained-scientific-innovation-cssi

[50] https://datascience.nih.gov/news/nih-awards-28-supplements-advance-software-tools-open-science

[51] https://wellcome.org/what-we-do/our-work/data-science-and-health-trustworthy-data-science

and further discussed in the Canadian RS Platforms section in this document.[52] The Canada Foundation for Innovation's Cyberinfrastructure Challenge 1 call for RS platforms (2016 and 2018) also brought significant direct funding to RS platforms.[53] Both of these programs have been discontinued through the transition of the Canadian RS mandate to the Alliance.

RS Sustainability

Most countries struggle with the balance between funding RS innovation and funding RS sustainability, with sustainability becoming critical as research relies more and more on RS. As Knowles et al. state:[54]

> *"... a key outstanding challenge is how we meaningfully evaluate requests for maintenance funding alongside new research proposals, because ultimately, funders' resources are constrained, and hard decisions will need to be made — what we need to ensure is that these are the right decisions."*

We borrow once again from the excellent ReSA survey on the current state of RS funding to list the rather short, but growing list of funders who are tackling this challenging issue. These include:

1. Chan Zuckerberg Initiative (CZI): Essential Open Source Software for Science;[55]
2. German Research Foundation (DFG):Research Software Sustainability;[56]
3. Netherland eScience Centre: eScience Centre Call for Sustainable Software;[57]
4. UKRI: Software for research communities;[58]
5. US National Science Foundation (NSF): CSSI Transition to sustainability;[59]

---

52 https://research-software-directory.org/organisations/digital-research-alliance-of-canada

53 https://www.innovation.ca/apply-manage-awards/past-competitions/2015-2017-cyberinfrastructure-initiative

54 Knowles et al., "We need to talk about the lack of investment in digital research infrastructure", *Nature Computational Science* 1, no. 3 (2021): 169-171, https://doi.org/10.1038/s43588-021-00048-5.

55 https://chanzuckerberg.com/eoss/

56 https://www.dfg.de/en/research_funding/programmes/infrastructure/lis/funding_opportunities/call_proposal_software/

57 https://www.esciencecenter.nl/calls-for-proposals/call-for-sustainable-software-2023-ss-2023/

58 https://www.ukri.org/opportunity/software-for-research-communities/

59 https://beta.nsf.gov/funding/opportunities/cyberinfrastructure-sustained-scientific-innovation-cssi

6. US NASA: Support for Open Source Tools, Frameworks, and Libraries[60]
7. US DOE: Seed Collaborations for Software Sustainability.[61]

In Canada, CANARIE has funded several rounds of RS development targeted at existing platforms with the goal of either broadening the use and/or extending the user base of the RS platform. These programs provide sustainability in the context of funding further development of existing platforms but did not target making RS platforms sustainable in the long term. In addition, CANARIE funded RDM platforms to expand their interoperability and increase adherence to the FAIR principles and international standards. The CFI *Cyberinfrastructure Challenge 1* funding provided Institutional Operating Funds (IOF) that could be used after the initial development funding to operate the platform. Importantly, by treating RS platforms as research infrastructure, the Infrastructure + Operating model that CFI typically applies to hardware/facilities was the first funding program in Canada to recognize that software platforms require operational support after initial development is completed.

RS Thematic Initiatives

Thematic research funding calls might fall into either RS Innovation or RS Sustainability categories, but are discussed separately because they are not broad, open calls but instead narrow specific calls to target specific problems. Such calls might look at specific domains such as the health sciences (e.g., CZI Single Cell Biology Data Insights)[62] or target specific domains and countries (e.g., CIHR/EU H2020 - International flagship collaboration with Canada for human data storage, integration and sharing to enable personalized medicine approaches)[63] or specific FAIR principles in specific countries (e.g., CHIST-ERA Open & Re-usable Research Data & Software)[64] or domains (e.g., DOE FAIR Data to Advance Artificial Intelligence for Science)[65].

Building Community

---

60 https://nspires.nasaprs.com/external/solicitations/summary!init.do?solId=%7B958CF134-D655-E512-B5AD-84501D14A0C1%7D&path=open

61 https://www.hpcwire.com/off-the-wire/doe-ascr-selects-seed-collaborations-for-software-sustainability/

62 https://chanzuckerberg.com/rfa/single-cell-data-insights/

63 https://cordis.europa.eu/programme/id/H2020_SC1-BHC-05-2018

64 https://www.chistera.eu/call-ord-announcement

65 https://www.energy.gov/articles/department-energy-announces-85-million-fair-data-advance-artificial-intelligence-science

Community-based funding initiatives are targeted at building the research community. These are often event-based programs that bring communities together in workshops or conferences or funding for the direct development of RS communities at a specific regional level.

In Canada the only RS focused event has been the CANARIE hosted Canadian Research Software Workshop (later called Canadian Research Software Conference, or CRSC).[66] These workshops were sponsored and run by CANARIE and CANARIE funded research projects received additional funding to send team members to this event. CANARIE not only sponsored the community event, but also directly funded community building via the Local Research Software Support (LRSS) funding calls. This funding program awarded local institutions for 2-3 years to build a local RS support team.[67]

The Compute Canada Federation (Federation) has historically hosted the Technical Experts of Compute Canada (TECC) meetings, which were internal events for Federation staff/partners to discuss national ARC issues and strategies. This event has been replaced by the Alliance's DRI Connect Event,[68] but again this is an internal Alliance event and is not open to the general community.

CANHEIT is the national meeting for higher education IT staff. [69] CANHEIT sometimes includes ARC and RDM components in their programs, but these events rarely focus on RS. In some cases TECC meetings and CANHEIT have been held together,[70] and these programs often have a stronger emphasis on ARC/RDM/RS programs.[71] Although no longer held, from the late 1980s to 2017, the Canadian International High Performance Computing Symposium (HPCS) was the preeminent Canadian venue to present research around ARC and the associated RS developments required to leverage new ARC technologies. [72] Finally, Canadian community based events such as Hackathons like HackSeq and Access exist as grassroots events put on

---

66 https://www.canarie.ca/event/crsc-2022/

67 ; https://www.canarie.ca/rdm-funding-call-recipients/; https://www.canarie.ca/to-equip-canadian-researchers-with-powerful-software-tools-and-expertise-canarie-awards-3-6m-to-six-institutions-to-fund-local-research-software-development-teams/

68 https://driconnect.alliancecan.ca

69 https://canheit2023.utoronto.ca/

70 https://canheit-tecc.sfu.ca/index.html

71 http://programme.exordo.com/canheit-tecc2018/delegates/presentation/86/

72 https://dblp.org/db/conf/hpcs/index.html; https://www.calculquebec.ca/en/nouvelle/registration-now-open-for-hpcs-2017/

by specific communities.[73] Currently there are no national level community RS events scheduled or planned in Canada.

Nationally and internationally, there are a number of similar models for establishing and developing RS communities. Many countries have developed either grass roots or funded RS societies. Starting in 2013 in the UK as the UK RSE Association, the Society of Research Software Engineers (SocRSE) currently has chapters at either the national or continental level, including Africa, Asia, Australia/New Zealand, Belgium, Germany, the Netherlands, Nordic, UK, and the USA.[74] Canada is a notable exception. These RSE societies grew out of the recognition that RSEs as individuals often have very limited career paths within institutions. In the UK where the RSE movement was born, the SocRSE has had a substantial impact on the establishment of formal RSE staff positions, the establishment of RSE groups at institutions, and a global movement around the importance of RSEs as a role in science. SocRSE offers financial support for events and initiatives which support establishing a research environment that recognises the vital role of software in research.[75]

A range of other community initiatives exists. Recently Schmidt Futures funded four international institutions for two years to build local research support teams for improving the quality of research and accelerating advancements, but also supporting longer term platforms and systems that encourage best practice in open scholarship.[76] ReSA is running a set of Funders Workshops to bring international funders together to discuss funding RS globally: the 2nd Workshop will be hosted by the Alliance in 2023 in Montreal, September 18-20.[77] Some domains, such as the Open Bioinformatics Foundation provides fellowships to attend conferences and workshops, aimed at increasing diverse participation at events promoting open scholarship practices.[78] The Sloan Foundation offers significant grants to institutions for establishing Open Software Program Offices to support Open Source RS.[79]

Some jurisdictions have taken the approach of rewarding individual's participation in the community, such as:

---

[73] https://www.hackseq.com/; https://accessconference.ca/past-conferences/2019-edmonton-ab/workshops-and-hackfest-2019/

[74] https://society-rse.org/

[75] https://society-rse.org/policy-for-socrse-events-and-initiatives-grant/

[76] https://www.schmidtfutures.com/our-work/virtual-institute-for-scientific-software/

[77] https://www.future-of-research-software.org/

[78] https://www.open-bio.org/event-awards/

[79] https://sloan.org/programs/digital-technology/ospo-loi

1. The Dutch eScience Fellowship Programme for community ambassadors around the use of research software;[80]
2. The Simons Foundation Scientific Software Research Faculty Award to support new research professor positions in existing US academic departments to be filled by scientific software-focused researchers;[81]
3. The Australian Research Data Commons Australian Museum Eureka Prize for Excellence in Research Software awarded for the development, maintenance, or extension of software that has enabled significant new scientific research;[82]
4. The European Master For High Performance Computing offers scholarships and awards for attending their training programs.[83]

EDI has been recognized as a challenge in the RS field, with the SocRSE establishing an EDIA Working Group to address some of these challenges and an active grassroots Diversity in RS community.[84] There are some funding programs that attempt to address this issue directly, such as the rOpenSci Champion Program, which "... *focuses on people who belong to groups that are historically and systematically excluded from the open software and research software communities*".[85]

Building Coordination

Coordination-based funding activities center around building and facilitating collaboration among stakeholders. These might be programs to build national infrastructures for supporting RS, national programs to develop RS HQP, and programs and initiatives to support and enable RS platforms.

Examples of programs that work across stakeholders to support the development of RS software can be found in several countries. Previously mentioned CANARIE's LRSS team call is such an example. Schmidt Futures' Virtual Institute for Scientific Software (VISS) funding program not only supports the launch of a network of scientific software engineering centers at

[80] https://www.esciencecenter.nl/calls-for-proposals/escience-center-fellowship-programme-2023/

[81] https://www.simonsfoundation.org/grant/scientific-software-research-faculty-award/?tab=rfa

[82] https://ardc.edu.au/article/enter-ardc-eureka-prize-for-research-software/

[83] https://eumaster4hpc.uni.lu/application/#elementor-toc__heading-anchor-2

[84] Chue Hong et al., "Understanding Equity, Diversity and Inclusion Challenges Within the Research Software Community", In *International Conference on Computational Science*, pp. 390-403. Cham: Springer International Publishing, 2021, https://arxiv.org/abs/2104.01712; https://society-rse.org/community/get-involved/edia-working-group/; https://diverse-rse.github.io/

[85] https://ropensci.org/blog/2022/09/22/launch-champions-program/

research universities, but more importantly ensures close collaboration among Schmidt Futures' network of Virtual Institutes. The Netherlands eScience Centre has developed a national team of RS HQP and these individuals are available for secondment to research projects for one to three years through a proposal/grant mechanism.[86] New Zealand has a Scientific Programmer Consultancy service, where RS HQP are embedded in a project for three to six months, with the goal of transferring RS expertise and skills into the research project team in a sustainable manner.[87] In Canada, regional ARC providers such as SHARCNET provide access to Programming Support through a competitive application process.[88]

Some countries/regions have funding programs that are designed to develop policy, approaches, and/or infrastructure to support RS. Horizon Europe has funded community-based projects that are targeted at developing community coherence and/or developing community guidelines that ensure integration of infrastructure, tools and services.[89]

Other countries have focussed on domain specific thematic development of RS platforms. The German NFDI program has funded domain specific consortia to develop research capabilities across many domains, with funding for consortia spanning the Social Sciences and the Humanities, Engineering Sciences, Life Sciences, and the Natural Sciences.[90] Australia's ARDC has taken a similar approach, historically co-funding research projects grouped across domains, with the ARDC recently moving towards leading large scale Program Level initiatives driven by their strategic goals, with recent domain specific Data Commons such as the Planet Research Data Commons (Earth and Environmental Science), People Research Data Commons (Human Health), and HASS+I Research Data Commons (Humanities and Social Sciences and Indigenous) receiving significant investments.[91]

RS Landscape in Canada

[86] https://www.esciencecenter.nl/calls-for-proposals/call-for-sustainable-software-2023-ss-2023/

[87] Pletzer et al., "How NeSI Helps Users Run Better and Faster on New Zealand's Supercomputing Platforms", *2017 IEEE 13th International Conference on e-Science (e-Science)*, Auckland, New Zealand, 2017, pp. 465-466, https://doi.org/10.1109/eScience.2017.73; https://www.nesi.org.nz/services/consultancy

[88] https://www.sharcnet.ca/my/research/programming

[89] https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/horizon-infra-2023-eosc-01-02

[90] https://www.nfdi.de/consortia/?lang=en

[91] https://ardc.edu.au/program/planet-research-data-commons/; https://ardc.edu.au/program/people-research-data-commons/; https://ardc.edu.au/program/hass-rdc-indigenous-research-capability/

As introduced in the Executive Summary, a coordinated and integrated approach to supporting RS as a pillar in DRI is new in Canada. In order to support RS as required in its mandate, it was necessary for the Alliance to perform a substantial amount of research on the current state of RS Canada. It is on this foundation that the 2025-2030 RS Strategy is based. The approach used to gain our current understanding of the RS landscape in Canada is presented below, and the resources cited are used extensively throughout this document.

The initial Alliance DRI Needs Assessment consultation process spanned all of the DRI pillars, and included a call for Community DRI Position Papers, a survey of stakeholders across the DRI ecosystem, as well as a set of Town Halls to share and validate findings to date.[92] This resulted in the publication of the *Researcher Needs Assessment* report in September of 2021. [93] This report contained extensive data on the critical nature of RS to researchers in Canada, with key findings from the overall report listed in Appendix A and a summary of key findings from the Community DRI Position Papers summarized in Appendix B.

In parallel to the DRI Needs Assessment, the Alliance RS Working Group, consisting of Alliance and stakeholder contributors, authored the first national RS position paper in the form of the *RS Current State Assessment.*[94] The report considered both the international and national RS ecosystem, with a comprehensive analysis of the challenges and opportunities for RS in Canada. This set the foundation for the analysis and recommendations presented in the 2025-2030 Strategy, with the Challenges and Opportunities section in this document based on those set out in the RS Current State Assessment.

In addition to these foundational reports, the RS Strategy Working Group undertook a number of initiatives to help fill gaps in our understanding of the RS landscape. Although the Needs Assessment survey included questions around RS, it was decided that a further survey would help provide a more complete understanding of how RS was used and developed in Canada. This survey was completed in June 2023 and the results are utilized throughout this report. [95] A set of focussed RS Town Halls were also held with senior team members from established RS Platforms, with the goal of understanding the needs and challenges faced by this critical part of the RS ecosystem. In order to ensure that value is delivered from the development of RS during the 2025-2030 mandate, a draft assessment of metrics for the success and value of RS was prepared (Appendix D). In order to quantify the current and desired future state of the RS

[92] https://alliancecan.ca/en/initiatives/canadian-digital-research-infrastructure-needs-assessment; https://alliancecan.ca/en/initiatives/position-paper-submissions

[93] https://alliancecan.ca/sites/default/files/2022-03/needsassessment_alliance_20220126.pdf

[94] Zhang et al., “Research Software Current State Assessment”. Zenodo, June 21, 2023. https://doi.org/10.5281/zenodo.8066970.

[95] Corrie et al., “Alliance Researcher Council Research Software Survey Report”, Zenodo, 2023, https://doi.org/10.5281/zenodo.8280449.

ecosystem in Canada, the RSSWG developed the Research Software Framework (RSoF). Based on the NIST Research Data Framework (RDAF) version 1.5, the RSoF framework uses the same Research Life Cycle model as RDaF with topics and sub-topics modified to encompass the research life cycle of RS.[96] The RSoF framework was used to assess the current state, desired state, and priority of a broad range of RS topics across the RS Life Cycle in Canada. The RSoF analysis is discussed briefly in the section The RSoF Framework applied in Canada and in detail in Appendix C.

These foundations are the basis for the RS 2025-2030 Strategy and we summarize the current state assessment of the RS Landscape in the following sections.

RS Funding in Canada

CANARIE's Research Software Program began in 2007 (and a similar RDM-focused program was launched in 2018) as a mechanism to fund development of research software tools that accelerate research discovery by enabling access to digital infrastructure and allowing researchers to focus on research instead of the supporting technology. [97] Since its inception, CANARIE's Research Software (RS) Program facilitated the development of RS tools that accelerate discovery and simplify access to DRI. It has awarded a total of $50.7M in funding, with 112 funded projects that cover many areas of academic research (180 disciplines), and created 145 RS tools that have been used by 1,191 research teams. CANARIE's RDM funding program allocated $4.5M to RDM platforms, funding a total of 13 projects and creating services and platforms that benefit 628 research teams. [98] Projects that received a portion of their funding through this program, and that are still active are listed in the Research Software Directory, curated by the Alliance.[99] In 2018, CANARIE launched a Local Research Software Support (LRSS) initiative.[100] Modeled partly on international efforts, funding was provided to local research software teams to develop, maintain, support, and evolve RS to equip Canadian researchers with RS tools and expertise within the institution.

The CANARIE RS Program was visionary and ahead of its time, but not without its challenges. It was the first, and one of the only funding programs, to provide direct RS platform development

96 Hanisch et al., "NIST Research Data Framework (RDaF): Version 1.5." (2023), https://www.nist.gov/publications/nist-research-data-framework-rdaf-version-15.

97 https://www.canarie.ca/software/

98 https://www.canarie.ca/program-news/rdm/

99 https://research-software-directory.org/organisations/digital-research-alliance-of-canada

100 https://www.canarie.ca/program-news/research-software/; https://www.canarie.ca/program-news/research-software/; https://www.canarie.ca/rdm-funding-call-recipients/

funding. It included innovation funding for new platforms, sustainability/reusability funding for existing platforms, interoperability funding for RDM platforms, and RS team development funding to catalyze the development of teams at institutions. The CANARIE programs resulted in the development of some of the top RS platforms in the world (See the Canadian RS Platforms section below) and led in the creation of the emerging RS community in Canada. Challenges with this program spanned legal, financial, and scale issues. The legal, administrative, and financial reporting requirements of the CANARIE program were substantial, requiring legal agreements between CANARIE and the institutions and significant financial reporting including full financial audits at the end of the project. Scale was also a challenge for the CANARIE program. Although the program recognized the importance of innovation, sustainability, team building, and research data management, it did not have the funding to support more than one of these at any given time. CANARIE ran consistent annual funding calls (see Graph below), but having to choose a single focus area (innovation vs sustainability, RS vs RDM) in any given year meant there was a gap in funding for other areas during that year.

The CANARIE RS funding mandate was transitioned to the Alliance in April 2022, *with no funding calls/awards for RS or RDM platforms being awarded since October 2020*. Given that the Alliance RS initiative in the 2023-2025 budget was not approved *there will be a five year gap in Canadian funding for RS, resulting in a sustainability crisis for RS platforms in Canada (See Figure 2)*. This strategy is intended to build on the strengths of the CANARIE and other Canadian RS funding programs, address the pain points and challenges, to fill in the existing RS funding gap, and to re-invigorate and re-build the RS community in Canada.

The Canada Foundation for Innovation (CFI) serves as an independent not-for-profit organization that funds research tools and infrastructure that help Canadian researchers conduct cutting-edge research and technology development, as well as to build and sustain a full spectrum of research infrastructure in Canada. In 2015 and 2017 CFI held its first RS funding program as part of its Cyberinfrastructure initiative.[101] Cyberinfrastructure Challenge 2 was targeted at renewing the ARC infrastructure in Canada while Cyberinfrastructure 1 was aimed at creating RS platforms that utilized the ARC infrastructure. This was CFI's first (and only) dedicated program in support of RS platforms. For the first time in Canada RS platforms that were developed through this program were treated as infrastructures, recognizing that they require ongoing funding to operate after they have been built and deployed. Projects receiving CFI Cyberinfrastructure Challenge 1 funding benefited from the CFI Institutional Operating Funding (IOF) which provided funding to operate the RS platform after the completion of the innovation required to build the platform.

[101] https://www.innovation.ca/awards/2015-and-2017-cyberinfrastructure-initiative

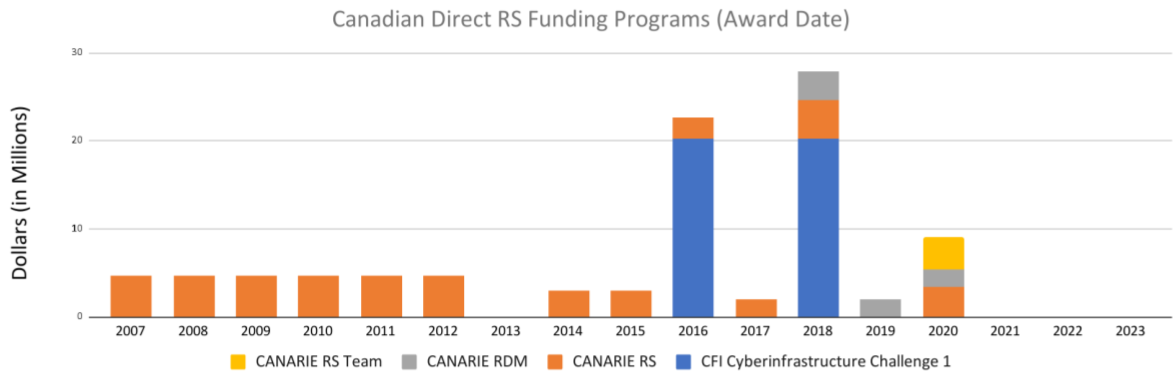

Figure 2. Direct RS funding programs from CANARIE and CFI with annual funding amounts based on funding program award date.

It has been noted that most of the national funding agencies are focused on funding research, and although many research projects develop RS there is little explicit funding for the development of research software or research infrastructure as the primary outcome of a research project. Most national research funding agencies and initiatives don't typically fund projects to explicitly develop RS, but instead do so in an indirect manner by funding projects that investigate a domain-specific research problem and, through that research, develop RS as an outcome of the project. Genome Canada/Quebec, while not focusing directly on RS, has supported the development of many bioinformatics tools and infrastructures in the support of other science-based genomics research projects.[102]

At the provincial/regional level, provincial governments are also key players in supporting SW innovation in higher education institutions. Although provincial/regional research funding agencies typically don't explicitly fund RS, there are several provincial and regional organizations that have an impact on the development and delivery of RS services. This support varies from province to province to territory, including direct funding (e.g., ACOA AIF) , (indirectly) matching funding, and a range of support programs.[103] There is also a recent trend to reimagine largely health-focused provincial funding agencies to support interdisciplinary and non-health focused research.

[102] https://www.genomecanada.ca/en/; https://www.genomequebec.com/; https://www.genomecanada.ca/en/software-peptide-identification-and-quantification-large-mass-spectrometry-data-using-data; https://www.genomecanada.ca/en/rapid-prediction-antimicrobial-resistance-metagenomic-samples-data-models-and-methods; https://www.genomecanada.ca/en/target-big-data-informed-software-design-cis-regulatory-regions-controlling-human-gene-expression; https://www.genomecanada.ca/en/cancer-genome-collaboratory.

[103] https://www.canada.ca/en/atlantic-canada-opportunities.html; https://www.canada.ca/en/atlantic-canada-opportunities/services/atlanticinnovationfund.html

Corporate research partnerships are also a hub for innovation and entrepreneurship, providing opportunities beyond the campus and bridging academia and industry, developing new technology and sourcing talent.[104] Many universities have created Technology Transfer offices to facilitate private sector partnerships, and spin-off activities from research, but most are focused on licensing and IP approaches. Open-Source Program Offices (OSPOs)[105] have emerged internationally, but they have yet to see adoption in Canada. By providing commercialization assistance, research-driven ideas are advanced; commercial research investment is leveraged; and students get opportunities to explore careers and gain work experience while earning a university degree. Some of these corporate research partnership engagements may also stem from the government funded or jointly funded programs, including some from NSERC, Mitacs, Ontario Centres of Excellence (OCE), SSHRC, etc.[106]

For a more comprehensive list of funding sources for RS development and use, see Appendix H in the Research Software Current State Assessment.[107]

RS Support in Canada

Besides serving as a RS funder, CANARIE played a critical role supporting and promoting RS best practices, advancing the Canadian RS ecosystem as well as cultivating the Canadian RS community. The CANARIE Research Software Portal (transitioned in 2022 to the Alliance's Research Software Portal) provides a central location for researchers and research software developers to discover software contributed by program participants, and by members of the greater research community.[108] Starting in 2018 CANARIE hosted an annual, discipline-agnostic event, the Research Software Workshop (more recently renamed Canadian Research Software Conference), which brought a community of RS professionals, developers and engineers from across the country together to exchange ideas and foster collaboration. CANARIE was also an active participant in the international research software community, by participating in International Research Software Engineering Leaders Workshop and served as a founding member of ReSA.[109] In 2017, 2018, and 2022, CANARIE collaborated with the UK Software Sustainability Institute (SSI) to conduct the annual Survey of Canadian Research Software

---

104 https://uwaterloo.ca/news/news/rogers-and-university-waterloo-partner-build-made-canada-5g

105 https://www.linuxfoundation.org/resources/open-source-guides/creating-an-open-source-program/

106 https://www.mitacs.ca/en

107 Zhang et al., "Research Software Current State Assessment". Zenodo, June 21, 2023. https://doi.org/10.5281/zenodo.8066970.

108 Researchsoftware.ca; https://research-software-directory.org/organisations/digital-research-alliance-of-canada

109 https://www.researchsoft.org/

Developers to better understand the needs of the Canadian research software community.[110] CANARIE completed the transfer of both its Research Software and Research Data Management programs to the Alliance on March 31, 2022 and the Alliance continues to participate in international initiatives such as ReSA. Given that the Alliance's RS initiative was not funded in the 2023-2025 mandate, there are no current plans to hold a dedicated RS Conference in 2023, although planning is underway for a DRI event, incorporating RS, in 2024.

As Canada's national ARC service provider, the Compute Canada Federation (CCF) was established to accelerate research and innovation by coordinating the deployment of state-of-the-art ARC systems, digital infrastructure, storage, and software solutions. These resources are freely accessible to Canadian academic research institutions, researchers, and their collaborators, playing an important role in helping maintain and develop the country's leadership role in international research. RS is of course a critical component of any ARC service. The CCF developed a comprehensive, unified, general-purpose software stack that is delivered via a distributed file system.[111] The RS applications and libraries in this stack fulfill their dependencies internally, without any requirements on the host system, allowing them to run on nearly every operating system, unlike the traditional approach which requires compiling, packaging and installing each software application for each particular version or flavor of operating system. This novel technique, in conjunction with delivering the software via a globally available filesystem, makes over 1200 RS applications available in directly executable form to end users anywhere in the world, without requiring them to install those applications on their systems. This work directly inspired the creation of the European Environment for Scientific Software Installations, a consortium pursuing this technique to improve the delivery of RS to scientific users in Europe.[112] This platform can also enable the expansion of ARC infrastructure into the community or commercial cloud through tools like Magic Castle, another Canadian innovation.[113] This CCF catalogue of available research SW tools, libraries and environments provides a common and unified suite of RS to researchers on all ARC systems.[114] With more than 200 experts, including research software experts, based at 37 partner institutions across the country, the CCF provided direct support to Canadian researchers by offering consultation, expert support, and regional training on a wide range of topics.[115] The CCF completed the transfer of its ARC services to the Alliance on March 31, 2022. The Alliance Federation

---

110 https://github.com/canarie/developer_survey_2017; https://github.com/canarie/developer_survey_2018; https://softwaresaved.github.io/international-survey-2022/; https://www.software.ac.uk/.

111 Boissonneault et al., "Providing a Unified Software Environment for Canada's National Advanced Computing Centers", In *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning)*, pp. 1-6. 2019, https://doi.org/10.1145/3332186.3332210.

112 https://www.eessi-hpc.org/

113 https://github.com/ComputeCanada/magic_castle

114 https://docs.alliancecan.ca/wiki/Available_software; https://docs.alliancecan.ca/wiki/Standard_software_environments

115 https://alliancecan.ca/en/services/advanced-research-computing/technical-support/training-calendar

continues to provide this broad expert support of ARC platforms and services, including significant RS support.

At the local level, many post-secondary institutions, research hospitals and/or affiliated institutes across Canada have some local RS service delivery by providing direct support to their researchers in terms of both infrastructure and access to RS tools, resources, research platforms, expertise, and training. Specifics for the scope of these services vary dramatically across institutions, from those that provide little or no services/support to well-developed or extensive suites of services for RS, dedicated centers for research computing, university-industry partnership, or RS resources that contribute to the provision of services at the regional and national levels. Researchers themselves are often actively engaged in creating and delivering RS inside and outside their research programs, particularly in research labs or through domain-specific communities of practice and professional societies. For many institutions, specific technology-focused researchers, or heads of research institutes, can be heavily involved in local RS efforts. As new forms (e.g., collaboration, publication, and methods for visualizing and analyzing data) of DRI emerge, the number of universities providing RS services and the scale of such investment required to support them is expected to increase significantly.

In a similar context, most national research labs have their own local IT and RS support. While government labs are generally not eligible for federal funding, they are able to participate in collaborative research efforts. Despite this internal funding and support, national labs do look to the Alliance and other national agencies for opportunities to share resources, typically by working with university-based collaborators. For example, the Canadian Advanced Network for Astronomical Research (CANFAR) consortium that runs on the Alliance ARC systems, is operated by the Canadian Astronomy Data Centre (CADC) at the Herzberg Astronomy & Astrophysics Research Centre (an NRC lab), while the PI is from Western University.[116]

Canadian RS Platforms

The above funding and support initiatives, although lacking a long term strategy and vision for RS in Canada, have been successful in facilitating the development and maintenance of a wide range of world class RS Platforms by Canadian research teams. Many of these platforms have national and international user communities. These platforms range from research planning platforms (DMP Assistant), national RDM platforms (FRDR, Lunaris, Borealis, etc.), through to a wide range of domain specific platforms.[117] Canada is a leader in Genomics RS Platforms, with the national response to COVID-19 resulting in the coordination of national Host and Virus

[116] https://www.canfar.net/en/

[117] https://assistant.portagenetwork.ca/; https://www.frdr-dfdr.ca/repo/; https://www.lunaris.ca/en; https://borealisdata.ca/

sequencing at the HostSeq and VirusSeq repositories.[118] Other genomics platforms such as the Cancer Collaboratory, CanDIG, GenAP, iReceptor, IRIDA, have both national and international user bases and have been widely cited in the literature (e.g., over 100 citations of the 2018 iReceptor paper according to Research Gate).[119] Canadian RS Platforms span the medical sciences (CBRAIN, LORIS), environmental sciences (Meridian, Motus, Ocean Networks Canada, PAVICS), space sciences (AuroraX, CANFAR), and humanities (LINCS, Canadian Writing Research Collaboratory).[120] As part of the CANARIE/Alliance transition, all platforms developed through past CANARIE funding are tabulated and documented for discovery and use by the Alliance on the RS Platform page and through a collaboration with the Netherlands eScience Center where the Alliance maintains this list in the international Research Software Directory.[121]

These platforms serve thousands of national and international users including iReceptor with 790 users (e.g., See Figure 3 - iReceptor user demographics), Motus with 1,800 users, Ocean Networks Canada with 84,000 users, and CANFAR with 8,300 data service users and with data distributed to over 40,000 unique IP numbers.[122] Platforms like Ocean Networks Canada are key for communicating science to the broader community.[123]

118 https://genomecanada.ca/challenge-areas/cancogen/hostseq/; https://virusseq.ca/

119 https://cancercollaboratory.org/; https://www.distributedgenomics.ca/; https://genap.ca/, GenAP will migrate to usegalaxy.ca in early 2024; https://gateway.ireceptor.org/; https://irida.ca/; Corrie et al, "iReceptor: a platform for querying and analyzing antibody/B-cell and T-cell receptor repertoire data across federated repositories", *Immunological reviews* 284, no. 1 (2018): 24-41, https://doi.org/10.1111/imr.12666.

120 https://cbrain.ca/; https://loris.ca/; https://meridian.cs.dal.ca/; https://motus.org/; https://www.oceannetworks.ca/; https://pavics.ouranos.ca/; https://aurorax.space/; https://www.canfar.net/en/; https://lincsproject.ca/; https://cwrc.ca/.

121 https://alliancecan.ca/en/services/research-software/canadian-research-software-platforms; https://research-software-directory.org/organisations/digital-research-alliance-of-canada

122 https://motus.org/; https://www.oceannetworks.ca; https://www.canfar.net/en/.

123 https://www.oceannetworks.ca/multimedia/live-cameras/

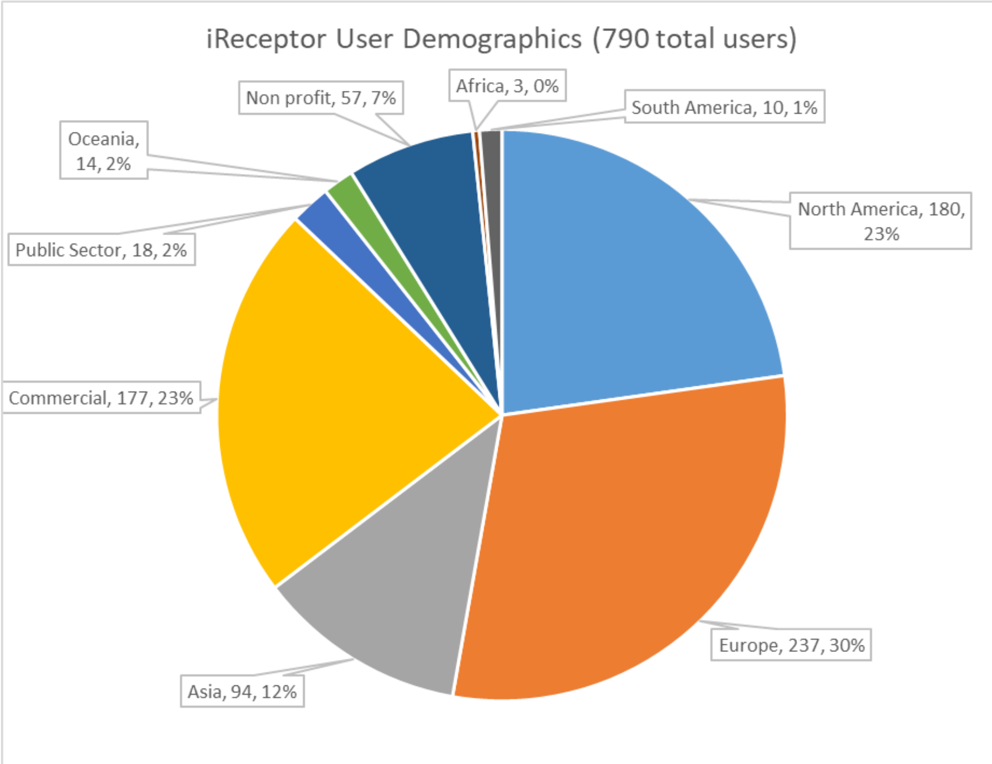

Figure 3. International demographics of iReceptor user accounts from 2014-2023

Many of these platforms have had a significant research impact, with extensive citations (e.g., over 100 citations of the iReceptor paper) for platform and data use/reuse.[124] Many of these platforms have been in production for several years (e.g., GenAP since 2013, iReceptor since 2014, and the Canadian Writing Research Collaboratory, or CWRC, since 2016), with sustainability being enabled by repeated success in acquiring innovation-focused grants rather than through RS programs that support sustainability (some later CANARIE grants being the exception). Over time, successful projects such as iReceptor and CWRC needed to rely on many grants (requiring many applications), some of extremely short duration (1.25 years). For instance, following initial investment from the CFI Leading Edge Fund, CWRC operations and maintenance funding has been spread across numerous sources (CFI IOF, JELF, and Cyberinfrastructure programs; CANARIE Research Software; SHARCNET; the CRC program; and modest contributions from research grants) requiring more than 12 grant applications focused on sustainability (rather than research) within a 7-year period. iReceptor (see Figure 4) has experienced both periods of no funding (and therefore no staff funding) and periods of substantial funding (with large, temporary teams); such variability is extremely challenging in terms of continuity of operations, efficiency, and lost expertise.

[124] Corrie et al, "iReceptor: a platform for querying and analyzing antibody/B-cell and T-cell receptor repertoire data across federated repositories", *Immunological reviews* 284, no. 1 (2018): 24-41, https://doi.org/10.1111/imr.12666.

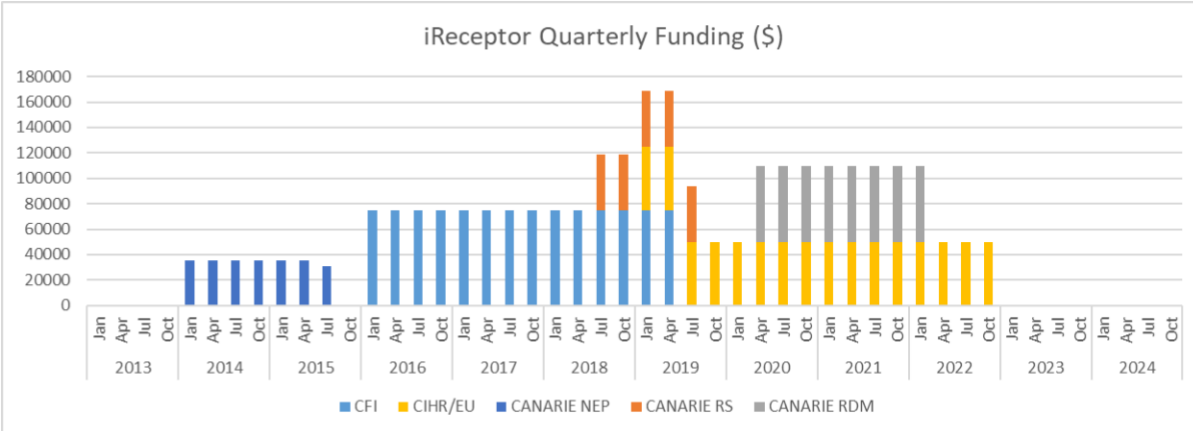

Figure 4. iReceptor quarterly project funding from project beginning (2014) to date, split across funding programs.

The Canadian funding environment therefore makes it very difficult to hire and sustain RSE teams and with the exception of large, multi-year programs such as the CFI Cyberinfrastructure and the CIHR/EU Horizon programs, no opportunity to plan beyond 12-18 months. This results in a large expenditure of researcher and staff time in writing grant applications and administering multiple grants simultaneously. In the case of iReceptor, failure to acquire any of the 5 funding awards shown would have meant a significant amount of time without any funding, threatening the ongoing viability of the platform. iReceptor's capability meant that the platform was in an ideal position to react quickly to the COVID-19 pandemic, with many citations for COVID-19 data reuse resulting from data curation efforts of COVID19 B-cell and T-cell data.[125] Only through the navigation of this complex funding cycle to sustain itself was iReceptor able to provide the service during the COVID-19 pandemic.

The fact that there has not been a direct RS funding call/award since 2020, with most 2020 awarded grants having completed at the end of 2022, has put some Canadian RS platforms that rely on these funding sources in a sustainability crisis. For example, the iReceptor project has maintained a project team of four people on average since the project started in 2014, with the same key developers members of the team since 2014. As of December 2022 the funding for these positions ended and these individuals have moved on to other positions in industry, government, or academia, resulting in a loss of 16 years worth of direct experience on this project. This leaves a single team member to maintain and develop the iReceptor platform until December 2023 when all Canadian funding will have ended, at which time the platform is at risk of being shut down.

Similarly, the Canadian funding environment makes it very difficult to maintain a career as a RSE in Canada. RS grants are typically short, and projects can often only offer confirmed

125 iReceptor citations: ResearchGate - Referenced 2023-08-23

positions to the end of these grants. Retaining HQP is therefore very challenging, with RS developers joining and leaving projects as funding changes. Recruiting, training, and then losing expertise in such a cycle is very costly and challenging for research projects. Research projects rely on attracting good candidates due to the academic work environment and interesting projects, but can not compete with industry salaries and benefit packages. Although some academic institutions provide RSE positions within university units, many RS developers rely on soft, cyclical grant funding. The RS Survey shows that most projects rely heavily on students and post-doctoral researchers for RS development, while at the same time these same roles are often either part-time or fixed term contracts (see Figure 5).

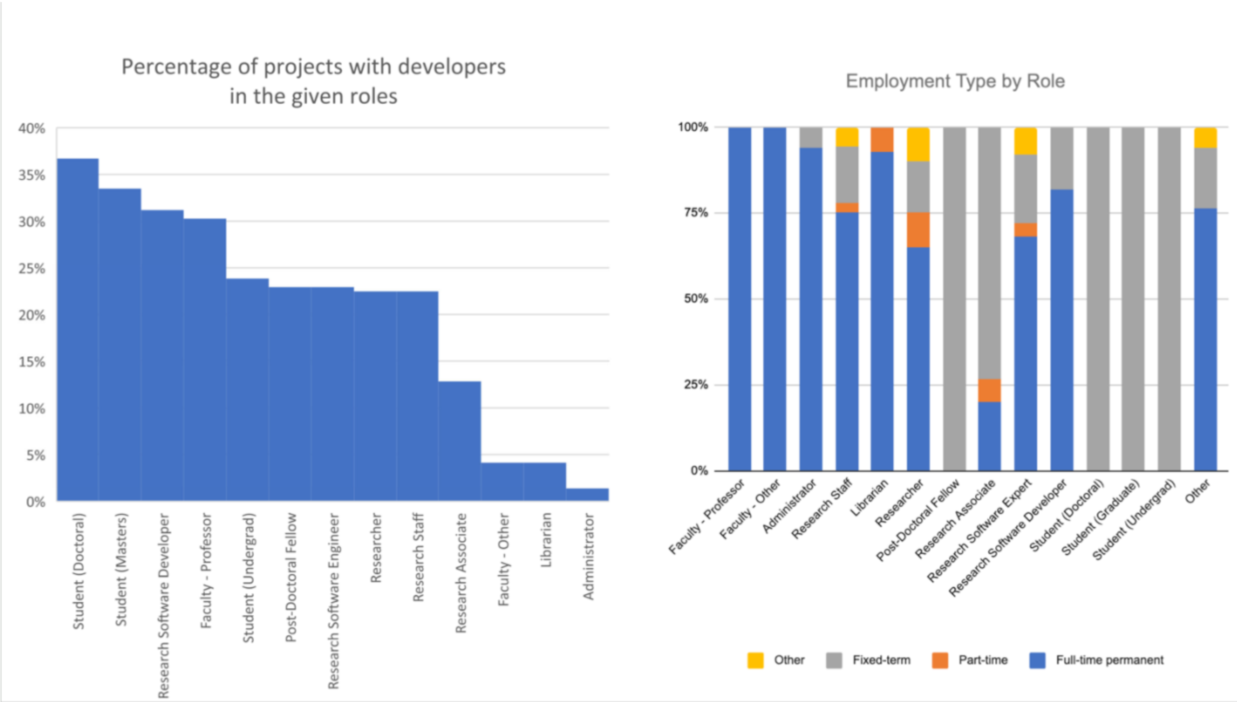

Figure 5. Left - Percentage of RS survey respondents that stated their projects had RS developers from specific roles in the DRI ecosystem. Right – Distribution of staffing levels (full-time, part-time, fixed term) for different roles on RS developers with fixed-term contracts (Students, Post-Doctoral Fellows, etc.)

When RS developers leave projects, they often leave the research sector completely for government or private sector positions, meaning that the DRI ecosystem will experience a cyclical sequence of finding, training, and losing highly skilled and valuable HQP. Although this cycle of training HQP in the research sector for the private sector does advance the private sector in Canada, it results in a very inefficient and fragile Canadian RS development capability. This in turn has a direct negative impact on the capabilities of researchers that rely on software development to advance their research. In particular, the lack of RS funding in Canada since 2020 has exacerbated this fundamental problem, with RS Platform PIs that attended the *Alliance RS PI Drop In Sessions* emphasizing the importance of longer term funding windows in helping to build highly skilled teams, the critical need to sustain these teams, and the large impact of losing key HQP with extensive domain knowledge. It was stated that there has already been significant loss of HQP from the DRI ecosystem in Canada due to the gap in funding since the end of the CANARIE and other targeted RS programs.

The RSoF Framework applied in Canada

This Research Software Framework (RSoF) was created by the Alliance RS Strategy Working Group in order to quantify the current state of the Canadian RS ecosystem, establish a desired future state for the 2025-2030, and to prioritize areas of importance. RSoF was derived from the NIST RDaF model (ver 1.5), and heavily modified to capture topics and subtopics of the Software Research Lifecycle (Envision, Plan, Develop/Acquire, Use, Share/Reuse/Preserve/Discard).[126] While the RDaF itself is still evolving, the WG felt that a framework-based approach would be a good fit for the Canadian RS ecosystem. Each of the 50 RS topics that were identified by the RSSWG were assigned both a current stated and a desired future state maturity level based on the Capability Maturity Model Integration (CMMI).[127] Topics were also assigned a priority from 1 to 5, establishing a mechanism to prioritize important topics over those that are less important (or possibly less urgent to be addressed in the current 2025-2030 mandate).

The RSoF is in one sense a conversational framework to facilitate dialog around the various components of research software, but can also serve as a matrix for inferring an organization's current readiness, or service level, in various contexts, as well as what the desired future state might be. The RSoF framework helps to link the research life cycle across RS and RDM (through the use of the RSoF and RDaF frameworks) as well as provides a mechanism to prioritize and focus on specific recommendations for the 2025-2030 Strategy. For the purposes of the 2025-30 mandate discussion, the current state and desired future state can then be used to determine the effort needed to “move the needle” by the desired amount, determine if that aspect is considered a priority, and what types of support programs might be a good fit.

For example, Figure 6 presents the current and future state of the Develop phase of the Research Software Lifecycle (as assessed by the RSSWG), which is split into 8 main topics that cover at a high level the development phase of RS. Each of these is assigned a current state maturity level (the blue line) as well as a desired future state for the end of the 2025-2030 mandate (the red line). As can be seen, the current state has been assessed as relatively immature currently (maturity ranging from 0-2), with a desire to move the state at the end of the 2030 mandate substantially in some cases (maturity ranging from 2.5 to 4.5). Similar assessments have been done for all stages of the RS Lifecycle. It is worth pointing out explicitly that none of the 50 topics within the RSoF assessment have a current state above a maturity

126 Hanisch et al., "NIST Research Data Framework (RDaF): Version 1.5." (2023), https://www.nist.gov/publications/nist-research-data-framework-rdaf-version-15.

127 https://cmmiinstitute.com/

level of 2, with this assessment confirming that RS is indeed immature in Canada. This assessment also establishes that although there is an enormous amount of potential growth for RS in Canada, achieving this potential will require a significant investment. The details of the RSoF assessments as well as more detail on the methods used to assign maturity and priority can be found in Appendix C.

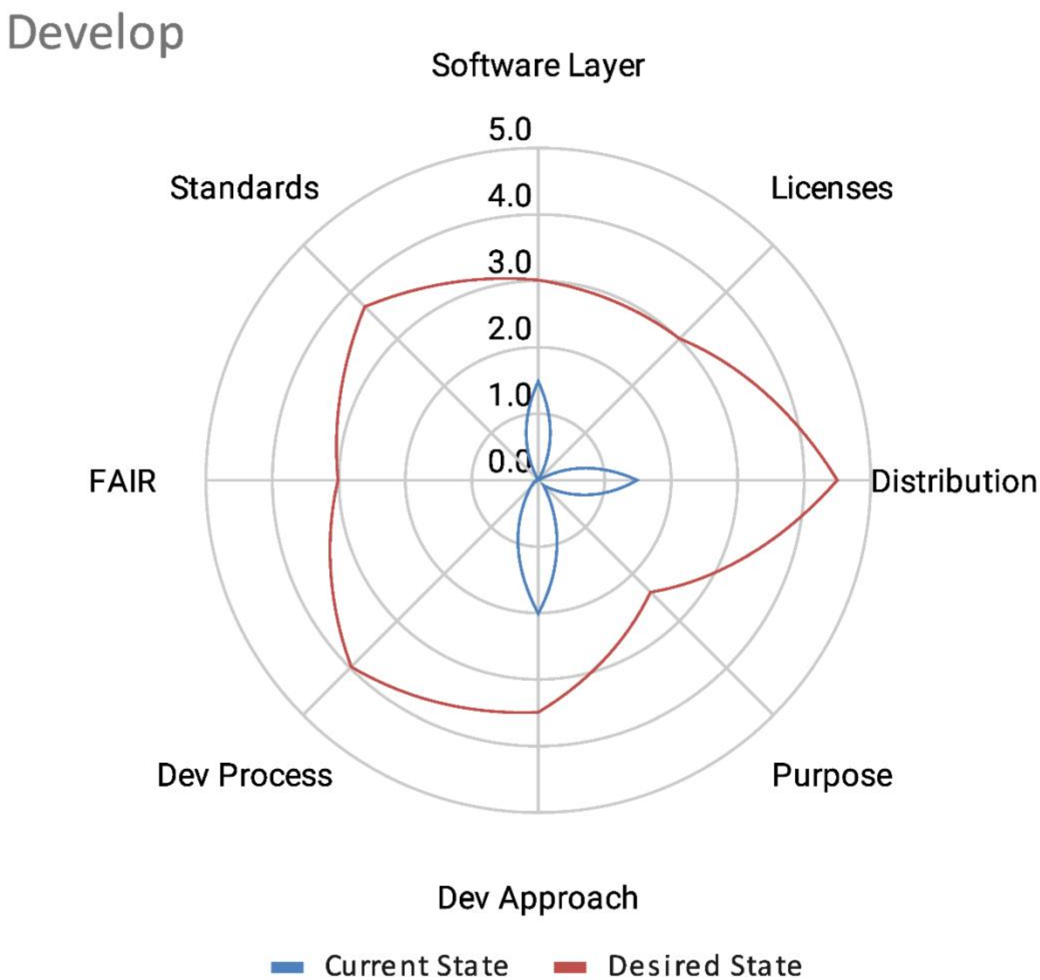

Figure 6. High-level summary of RSoF Topics for the Development stage of the Research Software Lifecycle. The blue line indicates the assessment of the current state of each RSoF Topic, while the red line indicates the desired state for each RSoF Topic.

Challenges and Opportunities

The RS ecosystem in Canada has established a significant RS capability, but this capability exists in a fragile and unstable funding environment. Released in late 2021, the Alliance *RS Current State Report* framed an understanding of the RS ecosystem in Canada: surveyed and summarized RS generally as an emerging field and as an area of professionalization, nationally and internationally, and documented strengths, challenges and opportunities.[128] This report, combined with the Alliance Needs Assessment, Alliance RS Survey, and Alliance RS PI Drop in sessions described above, have reiterated the many challenges facing the RS ecosystem in Canada while at the same time identifying significant opportunities for the Alliance to support and advance RS in Canada.

These challenges and opportunities are:

1. **Capability (Funding/Infrastructure/Research Data/Cybersecurity)**
    a. **There is insufficient support in targeted and sustainable RS funding**. Traditional research funding is fundamentally innovation-focused, leaving the need to sustain RS for the long-term up to the ability of PIs to describe their needs in an innovation context. This represents a significant gap and challenge for Canadian researchers, resulting in lost investment when RS is not maintained or generalized for reuse, and diminishes Canada's leadership role in the development of RS. Approaches to RS funding need to recognize the different types and phases of RS (RS innovation vs sustainability of existing RS platforms), and devise appropriate evaluative mechanisms, metrics, and funding streams for each. There is an opportunity for Canada and the Alliance to play a leading role internationally in supporting sustainable RS through adhering to and leading emerging initiatives such as the Amsterdam Declaration on Funding Research Software.[129]
    b. **Jurisdictional boundaries of funding sources prevent RS collaboration and innovation.** RS, like RDM, is multidimensional and international, as RS have no jurisdictional or other physical boundaries. Many Canadian RS platforms are international collaborations and/or have international user communities. RS funding, with few exceptions, has jurisdictional boundaries (e.g., Canadian funding calls are typically restricted to Canadian partners), limiting the ability for Canadian projects to collaborate internationally. There is a need for stakeholders (funders/governments) to work together to establish multilateral funding

[128] Zhang et al., "Research Software Current State Assessment". Zenodo, June 21, 2023. https://doi.org/10.5281/zenodo.8066970.

[129] Research Software Alliance. "Amsterdam Declaration on Funding Research Software Sustainability", Zenodo, August 1, 2023. https://doi.org/10.5281/zenodo.8325436, (see also https://www.future-of-research-software.org/, https://adore.software/).

programs (e.g., ReSA's Multi-lateral Funding Working Group) to support Canadian RS in providing a leading role internationally.[130]

c. **The RS ecosystem is varied and complex and evolving quickly.** RS is ubiquitous, and is used in its many forms across all infrastructure components (ARC, Cloud, GPUs, Storage). The DRI ecosystem, and in this context RS, needs to support all researchers across all technology platforms. RS development needs to track research advances and these advances are global and fast paced. It is necessary to help researchers navigate the DRI ecosystem in order to help them find a RS solution for their research problem that is sustainable over time.
d. **RS is not widely disseminated or shared and not readily discoverable, inhibiting research transparency, reproducibility, and verification**. For science to be open and reproducible researchers need to acknowledge and cite RS when it is used in their research. At the same time RS developers need to deposit and publish their code so it can be cited. National services involving training for RS publishing/citation as well as platforms for cataloging, indexing, preservation, and curation of RS will be critical to developing an innovative and sustainable RS ecosystem in Canada.
e. **There is no coordinated approach to address research security and integrity and protect Canadian interests**. RS teams are typically ill-positioned to understand or to respond to the proliferation of cybersecurity threats, as these threats require specific skill sets and expertise in best practices for secure software development, as well as an understanding of data privacy policy. It is essential to include RS development as a fundamental component of a robust national research cybersecurity strategy and associated response, such as CANARIE's and the Alliance's own cybersecurity initiatives.
f. **There should be a balance between intellectual property (IP) management and developing a culture in which sharing information and research outputs is advocated.** This culture is best fostered by individual research communities, as norms vary by discipline. The ultimate goal is for research outputs to be appropriately protected, while fostering an open and collaborative approach to research that advances Canada's world-leading research and development.

2. **Community (Events/Training & Support)**
   a. **In comparison to many other countries, Canada lags behind in support of an effective, cohesive, and coordinated RS development community.** RS in Canada is typically viewed as a byproduct of domain-specific research and is not recognized as "first class" output or a core component of research. By contrast, other countries such as the UK, Australia, US, Netherlands, Germany, (and the EU in General) have both grass roots RS community organizations and/or invested large sums in a focused effort to build robust RS development

[130] https://www.researchsoft.org/taskforces/

communities across disciplines. The Alliance RS Survey indicates that only 3% of participants are currently part of an RS association and approximately 70% indicate that they would or might join such an association if it existed in Canada.

b. **There is a lack of SW development skills at all stages of the education and research pipeline.** There is a shortage of training in RS, with a particular need for RS training in those domains that are newer to the use of DRI in support of their research. The Alliance RS Survey indicates that the majority of respondents are self taught, few have received RS training from local/regional services such as the Alliance, and 72% would be interested if such training was offered. HQP skilled in RS, and in particular ensuring that those HQP stay in the research ecosystem, will have a long term effect on the efficiency and productivity of research in Canada.

c. **The lack of diversity in the RS community suggests we need a more effective approach to Equity, Diversity, and Inclusion (EDI) in the RS context**. The lack of diversity in the RS community suggests a more effective approach to Equity, Diversity, and Inclusion (EDI) in the RS context. The Alliance RS Survey shows that 52% of respondents identify as male while 30% identify as female, 13% identify as being from a racialized group, and only 1% identify as being Indigenous. All programs/services developed to support RS (community, events, training) should be scrutinized to ensure that appropriate measures are taken to maximize EDI.

3. Coordination (Policy/Governance)

a. **Canada does not incentivize RS and metrics for funding, reward and recognition, and career progression do not apply to RS outputs.** There is an opportunity for all stakeholders to enact policies, programs, and assessments that recognize the important role RS plays alongside other research outputs in achieving research outcomes in all disciplines. Appropriate metrics for RS activity would facilitate the culture change needed to realize the promise of open scholarship.

b. **Canada currently does not have well-developed policies, standards, and protocols to ensure adherence to international RS best practices, standards, and interoperability**. While few would argue that today's research enterprise is multidisciplinary and global, fewer have an appreciation for how we can ensure the global interoperability of research infrastructures to provide robust support for all forms of research. It is necessary to support national RS development efforts so that these efforts are part of the global RS fabric.

c. **HQP (highly qualified personnel; including researchers, students, postdocs, and research support staff) that develop RS need to be recognized for their contribution to the scientific process.** "Research software engineer" (RSE) is a recently minted job title that recognizes the extent to which research efficiency and outcomes are advanced by having HQP with RS skills on research teams.[131] Although there are nascent RSE communities in most regions, Canada requires focused efforts to establish both a national RSE community and a stable career path for RSEs within their host institutions and beyond. A robust recruitment and retention strategy for RS HQP, and especially RSEs, is needed.

d. **Federal and provincial research organizations need to develop a strategy for working more collaboratively with higher education and across governments.** Although national coordinated leadership for RS is emerging, efforts to crystallize RS communities have been hampered by a lack of adequate funding and a formal mandate. Without coordination of investment in the context of RS, it is difficult to develop the shared policies, processes, protocols, best practices, and standards that are so essential.

[131] Even though we are using this commonly accepted term in this document, in Canada RSEs are not necessarily accredited engineers.

A Look Forward: Future RS Ecosystem 2023-2030

Vision

To build a world-class, equitable, and sustainable research software ecosystem in Canada through advancing capability, community, and coordination, delivering research software that spans the DRI ecosystem and enables advances in research and innovation.

Goals

1. **Build Capability** by ensuring that Canada has the capability/capacity to **develop** and **sustain** world-class RS Tools and RS Platforms.
2. **Build Community** by developing an active and skilled Canadian community (**HQP**) around the **use** and **development** of RS.
3. **Coordinate** with stakeholders to provide **governance** of the RS **ecosystem** and **establish** Canadian RS **policy.**

Recommendations

Given the nature of stakeholder feedback, the results of the RS survey, and the current/desired states and associated RSoF assessment, we recommend that the Alliance focus on the following initiatives in the 2025-30 mandate. These recommendations are centered on the role of the Alliance, and the forms of support that it can provide according to its Contribution Agreement and mandate. Critical to success for RS in the Alliance's 2025-30 mandate is to make recommendations that support and enhance national RS resources and services, that intersect with regional and institutional efforts in a synchronized and effective way, that advance the adoption of (inter)national RS best practices, and that orchestrate and overlap with the other DRI "pillars". The goal of these recommendations is to meet researcher, developer, institutional and disciplinary needs, improve the visibility, impact, and reusability of RS, and ultimately create a coherent, coordinated and sustainable DRI ecosystem.

A detailed list of recommendations, spanning the goals of advancing Capability, Community, and Coordination of RS in Canada, are given below (see Figure 7). Each of these recommendations is associated with one or more RSoF Lifecycle Stages/Topics along with a value for the topics Current State (CS), Desired State (DS), and Priority (from 1 to 5, 5 being highest). Across the stages of the Research Lifecycle, a total of 16 of the 50 RSoF topics have been assigned Priority 5. Each recommendation is associated with at least one Priority 5 RSoF topic and all Priority 5 topics in the RSoF table are listed in the recommendations. Viewing the RS ecosystem through the RSoF lens allowed us to focus on areas where attention could be paid across the complete research lifecycle and spanning ARC and RDM. More detail on all 50 of the RSoF topics is available in Appendix C.

Figure 7. Recommendation categories, grouped across the goals of building RS Capability, Community, and Coordination.

Capability

Funding

1. **Provide RS innovation funding**. Develop and launch a merit based, innovation focused funding program for new and existing RS platforms such that Canada can remain a leader in the development of such platforms.
    - Envision - Sustainability: CS = 0.3, DS = 4.0, Priority = 5
    - Plan - Financial: CS = 0.1, DS = 4.0, Priority = 5
2. **Provide RS sustainability funding**. Develop and launch RS sustainability programs (funding) and provide HQP support (FTEs) for Canadian developed national and international RS platforms.
    - Envision - Sustainability: CS = 0.3, DS = 4.0, Priority = 5
    - Plan - Financial: CS = 0.1, DS = 4.0, Priority = 5
    - Preserve/Discard - Sustainability: CS = 0.3, DS = 3.0, Priority = 5
3. **Deliver regular and consistent funding calls under the RS innovation and sustainability program.** Enable a stable and expanding RS platform ecosystem by providing a well structured and regular funding cadence, including the consideration of how funding should be managed across mandate renewals.
    - Envision - Sustainability: CS = 0.3, DS = 4.0, Priority = 5
    - Plan - Financial: CS = 0.1, DS = 4.0, Priority = 5
4. **Participate in RS funding activities nationally and internationally.** Collaborate with primary funding agencies to develop strategies and funding programs that recognize RS as a first-class output of research and help establish consistent and sustainable funding approaches to RS.
    - Envision - Sustainability: CS = 0.7, DS = 4.0, Priority = 5
    - Plan - Financial: CS = 0.1, DS = 4.0, Priority = 5
    - Preserve/Discard - Sustainability: CS = 0.3, DS = 3.0, Priority = 5

Infrastructure

5. **Develop a Service Catalogue for RS platforms, tools and services.** Work with stakeholders to develop a comprehensive list of RS platforms that have been developed within Canada and provide resources and tools for sharing this information.
    - Share/Reuse - Publishing: CS = 0.7, DS = 2.5, Priority = 5
    - Share/Reuse - Dissemination: CS = 0.4, DS = 2.6, Priority = 5
    - Share/Reuse - Attribution: CS = 0.4, DS = 3.1, Priority = 5
6. **Create a middleware RS development capability to support RS platforms**. Ensure that the ability to develop and support middleware (authentication, job management, resource discovery, data discovery) for DRI resources in Canada (ARC/RDM) is appropriately resourced (FTEs) such that RS projects do not have to develop their own solutions (e.g., cyber-security) for which they are unlikely to have the appropriate skill set.
    - RS Develop - Standards: CS = 0, DS - 3.6, Priority = 5
7. **Improve support of users in the managing and use of RS platforms**. Ensure that researchers that are using RS platforms that span the DRI pillars are adequately supported. Provide adequate resources (FTEs) to support the use of RS platforms at a national scale. Support for researchers developing such platforms (DevOps, monitoring) as well as support for installing and operating existing platforms, both Canadian (e.g., FRDR) and externally developed (e.g., Galaxy), that have a broad applicability to the research community.
    - RS Use - Management: CS = 1.7, DS = 4.0, Priority = 5

Research Data

8. **Provide funding and expertise to ensure RS platforms manage data according to appropriate regional, national, and international legislative and regulatory frameworks.** Ensure that the RS/RDM intersection has adequately resourced expertise (FTEs) such that RS developers have access to appropriate training and expertise around the principles of RDM and FAIR data.
    - RS Develop - FAIR: CS = 0.0, DS = 3.0, Priority = 5

Cybersecurity

9. **Provide funding and expertise to ensure Canadian RS tools, platforms, and services are being developed and disseminated in an appropriately secure manner.** Ensure that cybersecurity is adequately resourced (FTEs) such that RS platforms have the ability to find resources within the DRI ecosystem to advise/consult on cybersecurity best practices.
    - Envision - Governance: Legal: CS = 1.0, DS = 2.9, Priority = 5
    - RS Develop - Standards: CS = 0, DS - 3.6, Priority = 5

Community

Events

10. **Reinvigorate the Canadian Research Software Experts (RSE) community**. Work with stakeholders and provide resources (FTEs, funding) to re-establish Canadian national RS events, facilitate community based RS activities (e.g., national slack channel for RSEs, Canadian RSE Society), and increase Canadian participation in international RS organizations.
    - Envision - Community: CS = 0.3, DS = 2.8, Priority = 5

Training & Support

11. **Ensure that training for RS development is explicitly included in the national training framework and resourced appropriately.** Work with internal (ARC and RDM training initiatives) and external (institutional, regional, national and international) training groups and content developers to ensure that RS development (software development best practices) training is adequately represented (content) and resourced (FTEs), including the important topics of the FAIR principles of RS (FAIR4RS) and the importance of RS reproducibility and reuse.
    - Plan - Training: CS = 1.5, DS = 3, Priority = 5
    - RS Develop - FAIR: CS = 0.0, DS = 3.0, Priority = 5
    - RS Develop - Dev Process: CS = 0.2, DS = 4.0, Priority = 5
    - RS Use - Management: CS = 1.7, DS = 4.0, Priority = 5
    - Share/Reuse - Publishing: CS = 0.7, DS = 2.5, Priority = 5
    - Share/Reuse - Dissemination: CS = 0.4, DS = 2.6, Priority = 5
    - Share/Reuse - Attribution: CS = 0.4, DS = 3.1, Priority = 5
12. **Increase the capability of training and support of the use of RS in the cloud.** Ensure that training and support for emerging Platform-as-a-Service offerings from the ARC/Cloud team are developed (content) and resourced (FTEs) appropriately. Provide tools, expertise and support for code development and management including Continuous Integration and Continuous Deployment (CI/CD) pipelines. Provide training (content) and resources (FTEs) that can advise researchers on how best to utilize the DRI ecosystem for their research.
    - RS Use - Management: CS = 1.7, DS = 4.0, Priority = 5
    - RS Develop = Dev Process: CS = 0.2, DS = 4.0, Priority = 5
13. **Expand existing RS consulting capabilities.** Collaborate with institutions (e.g., McMaster) and regional consortia (e.g., SharcNet) to provide RS consulting for in-depth support and training of RS teams. Provide resources for hiring a core set of FTEs that can be assigned to projects to help researchers develop/optimize code and improve their RS development skills.[132]
    - RS Develop - Dev Process: CS = 0.2, DS = 4.0, Priority = 5

[132] https://research.mcmaster.ca/home/support-for-researchers/research-resources/research-software-development/; https://www.sharcnet.ca/my/research/programming.

14. **Develop a mechanism and offer regular calls for projects to apply for access to RS consulting.** Collaborate with stakeholders to develop a national program for RS projects to apply for access to RS consulting to improve both the quality of RS software and the RS development skills within their teams.
    - RS Develop - Dev Process: CS = 0.2, DS = 4.0, Priority = 5

Coordination

Policy

15. **Reflect the RS community's needs in policy.** Facilitate the coordinated development of national funder policies regarding RS management planning, RS citation and RS reuse.
    - Envision - Governance: Strategic: CS = 0.8, DS = 3.2, Priority = 5
    - Envision - Governance: Legal: CS = 1.0, DS = 2.9, Priority = 5
16. **Work towards establishing RS as a "first-class" output of research.** Work on policy development with national and international stakeholders.
    - Envision - Governance: Strategic: CS = 0.8, DS = 3.2, Priority = 5
    - Envision - Culture: CS = 0.4, DS = 2.2, Priority = 5

Governance

17. **Ensure RS programs are undertaken with equity, diversity, inclusivity, and accessibility as core components**. Make EDIA considerations part of all funding calls, services and other components of the national DRI.
    - Envision - Governance: Strategic: CS = 0.8, DS = 3.2, Priority = 5
18. **Support international and national RS organizations**. Provide funding for, and participate in, governance forums.
    - Envision - Community: CS = 0.3, DS = 2.8, Priority = 5
19. **Work with stakeholders to establish robust career paths for RSEs.** Collaborate with stakeholders to establish stable career paths within institutional departments (IT Services, VPRs office, Library) wherever possible.
    - Envision - Sustainability: CS = 0.3, DS = 4.0, Priority = 5
20. **Demonstrate improvement through performance indicators and metrics**. Ensure that RS programs are progressing through the assignment of key performance indicators and the establishment of metrics to measure improvement. Ensure that the monitoring of programs is appropriately resourced (FTEs).
    - Plan - Assessment: CS = 0.5, DS = 4.0, Priority 5

Tracking Progress and Impact of RS in Canada

Gathering metrics to measure the impacts of research software is critical, but it can be complex and will vary depending on the types of applications, their implementation, and the domain in which they are applied etc.[133] The high-level list below was derived from a more detailed list of metrics (see Appendix D) used in various national and international contexts and is provided here to facilitate the discussion of appropriate metrics and KPIs around the success of specific RS platforms and tools.[134] Given that the citation of software in the research literature is relatively immature, AI tools intended to extract mentions of research software tools and platforms in publications can be used to derive an impact measure as well. Although difficult to list categorically in order of importance, metrics that attempt to measure research impact are listed first.

Possible metrics fall into the following categories:

1. **Citations**: e.g., RS paper citations (for one or more of RS use, RS Platform used for data storage, RS Platform source for data reuse), RS software mentions (RS noted in paper, but no formal citation), RS software citations (github release citation).
2. **Data metrics:** e.g., number of repositories, number of depositors/data stewards, number of data sets (research study level aggregation), number of entities (where studies may have multiple entities), number of records (entities might have many records), number of downloads.
3. **User metrics**: e.g., total users (lifetime of platform), unique active users (fixed period), number of logins/visits, number logins/unique users, new users, geographic distribution of users.
4. **Project/Funding metrics:** e.g., number of years under development, number of years in production (used by groups other than self), number of successful RS funding proposals, number of unsuccessful RS funding proposals, total budget over project duration, minimum annual budget over project duration, staffing levels.
5. **Operations**: e.g., storage costs, compute costs, total hardware costs, cost/dataset, cost/download, cost/analysis, staff costs.
6. **Web metrics**: e.g., visits/sessions, page view, hits/requests, downloads, unique users/IP address, geographic location, uptime.

---

[133] https://rse.shef.ac.uk/blog/2022-10-13-quality-value-research-software/

[134] Stockinger et al., “Plan for collation of metrics and quality data at the ELIXIR Hub”, Zenodo, March 8, 2018, https://doi.org/10.5281/zenodo.1194122; Metrics for Data Repositories and Knowledgebases: Working Group Report, https://datascience.nih.gov/news/metrics-for-data-repositories-and-knowledgebases-working-group-report

It is important to note that these are not metrics and KPIs that measure the success of the RS Strategy but instead are metrics that measure the success of individual RS platforms and tools. When the recommendations put forward in the RS Strategy are operationalized, additional KPIs that are measurable and quantifiable should be assigned for each.

Risks

The RS Strategy Working Group (RSSWG) has assessed some of the risks related to the 2025-2030 RS Strategy. It is recommended that the Alliance should tabulate these risks as part of the Alliance's risk register, track changes to the Strategy and how they impact these risks, and provide risk mitigation as required as part of an operational plan to implement the Recommendations. The risks listed below (Table 1) are specifically listed here because the RSSWG felt that these risks are specifically important to RS. The RSSWG did not feel that it was their role to propose mitigation strategies for these risks, as it is the Alliance's responsibility to determine risk mitigation strategies for Alliance initiatives as part of operational planning.

Where appropriate, risk titles and descriptions were taken from existing Alliance Risk documents, with possible modifications as pertinent to RS specifically. The original source for each risk is stated in the "Source" column. The source for these risks are drawn from existing Alliance Risk Registries as noted.

Table 1. Risk Register, based on the Risk Regsiter in the Alliance 2023-2024 Corporate Plan, with risks specific to RS noted explicitly.

The Alliance's risk register	RS Specific Risk	Source
Continuity and quality of service: The transition of ARC, RDM and RS to the Alliance combined with the continuity of current operations may be complex to execute without impacting quality of service. The transition to a production-type environment will require Service Level Agreements and KPIs (Key Performance Indicators) to assess performance. Current resources may not have the skills and/or experience to manage this transition. In addition, retention of Alliance DRI professionals may prove difficult the longer the transition takes.	**Continuity and quality of service: Lack of continuity in programs may damage the RS community as retention of skilled professionals may prove difficult the longer the transition takes. RS platforms may be lost due to gaps in funding and support.**	2023-2024 Corporate Plan[135]
Expectations and timing: Initial projects delivery timelines may lag given the complexity of the subject matter, the aggressive schedule and limited resource availability. This could affect the Alliance's ability to fully meet community expectations. There is also a risk of fragmentation among the DRI community—including competing funding proposals and regional and	**Expectations and timing: There is a risk of damage to the RS community, loss of RS Platforms, and a risk of loss of HQP from the DRI ecosystem due to delays in RS programs.**	2023-2024 Corporate Plan

[135] https://alliancecan.ca/sites/default/files/2023-04/2023-24%20Corporate%20Plan.pdf

discipline specific approaches rather than a unified national approach—if the Alliance cannot meet expectations.		
Funding: Inability to finalize the Contribution Agreement; inability to generate match funding with provinces and other sources; and the rejection of the National Service Delivery and Funding Model, and Strategic Plan.	**Funding**: **Inability to run programs due to the rejection of funding initiatives.**	2023-2024 Corporate Plan
Reputation: Failure to manage stakeholder expectations, failure to serve the research community; the Alliance Contribution Agreement not met.	**Reputation**: F**ailure to deliver on mandate to fund, support, and coordinate RS**.	2023-2024 Corporate Plan
Domain focus: Although ARC is more complex and involves more resources (Alliance DRI professionals and budget), transition planning for ARC should not be to the detriment of RDM and RS.	**Under resourcing RS - Domain focus**: **In particular, RS is less mature than ARC and RDM and is therefore RS is particularly susceptible to this risk.**	2023-2024 Corporate Plan

Proposal for 2023-25

With the fact that there was no RS funding from 2021-2023, plus the rejection of the original Alliance RS proposal to ISED for the 2023-25 mandate, some specific high-level risks and a recommended approach to mitigate those risks are listed below.

Risks include:

1. Loss of HQP through inconsistent and erratic funding of RS in Canada.
2. Loss of valuable RS Platforms through lack of approach to sustainable RS funding.
3. Loss of confidence in the Alliance as a DRI provider.
4. Lack of support from ISED for a comprehensive approach.
5. Lack of funding in the transition from CANARIE/CFI funding to the 2025-30 implementation results in a diminished RS community in Canada.

Given that the original Alliance RS proposal to ISED for the 2023-25 mandate was not approved, the WG felt that there is an urgent need to facilitate the work of existing Canadian RS teams in a way that intersects strongly with the Strategy outlined here.

1. Deploy a 2024-25 call that speaks to the emerging strategy, but provides some level of support for existing RS platforms and teams.
    a. Description: During the 2023-24 timeline, the Alliance will be working on the details of the NSDM, including how to review and designate RS platforms as national in scope. An initial version of that work will be used to identify specific national/international platforms in the Canadian community to receive transition funding. This funding will facilitate ongoing enhancement as needed, and also the integration of these platforms into the Alliance's national service framework.
    b. Funding
        i. Directed Calls
        ii. $5 million
    c. Model
        i. National / International
    d. Start/End
        i. 2023-25

Appendix A: Alliance Needs Assessment

The Alliance's Needs Assessment report was published in September 2021, reflecting the highlights from the 2021 Needs Assessment process, including a Survey, a series of Town Halls, and 100+ Position papers.[136] The final document highlighted a number of high-level items of relevance to the RS ecosystem.

1. Create a national catalogue of DRI resources that is accessible and available to all researchers in Canada, allowing them to rapidly identify the right solution for their needs.
2. Provide equitable funding for institutional support, especially additional streams of funding for smaller institutions, facilitating the hiring of Professional Support Staff and strengthening local service provision.
3. Develop Open Scholarship mandates and policies in collaboration and coordination with national stakeholders to ensure alignment between compliance requirements and service provision across all levels.
4. Adopt shared standards or policies for interoperability, recognizing that the national DRI should be capable of integration with international systems and enable international cooperation.
5. Provide a federated secure and standardized solution to address multiple privacy regulations is needed to cope with the prevalent challenges in the interoperability of datasets (e.g., data linkage, common ontologies) and platforms, particularly in managing sensitive data in the health and life sciences, as well as in disciplines working with qualitative data, due to frequent lack of widely adopted standards.
6. Adopt and promote the use of free and open-source hardware and software.
7. Provide professional services for the development of code, algorithms, and pipelines.

This graph (Figure 8) from the Needs Assessment Survey is instructive: of the top 10 activities that are part of a researcher's workflows, 6 are directly related to RS.

[136] https://alliancecan.ca/sites/default/files/2022-03/needsassessment_alliance_20220126.pdf; https://alliancecan.ca/en/initiatives/canadian-digital-research-infrastructure-needs-assessment

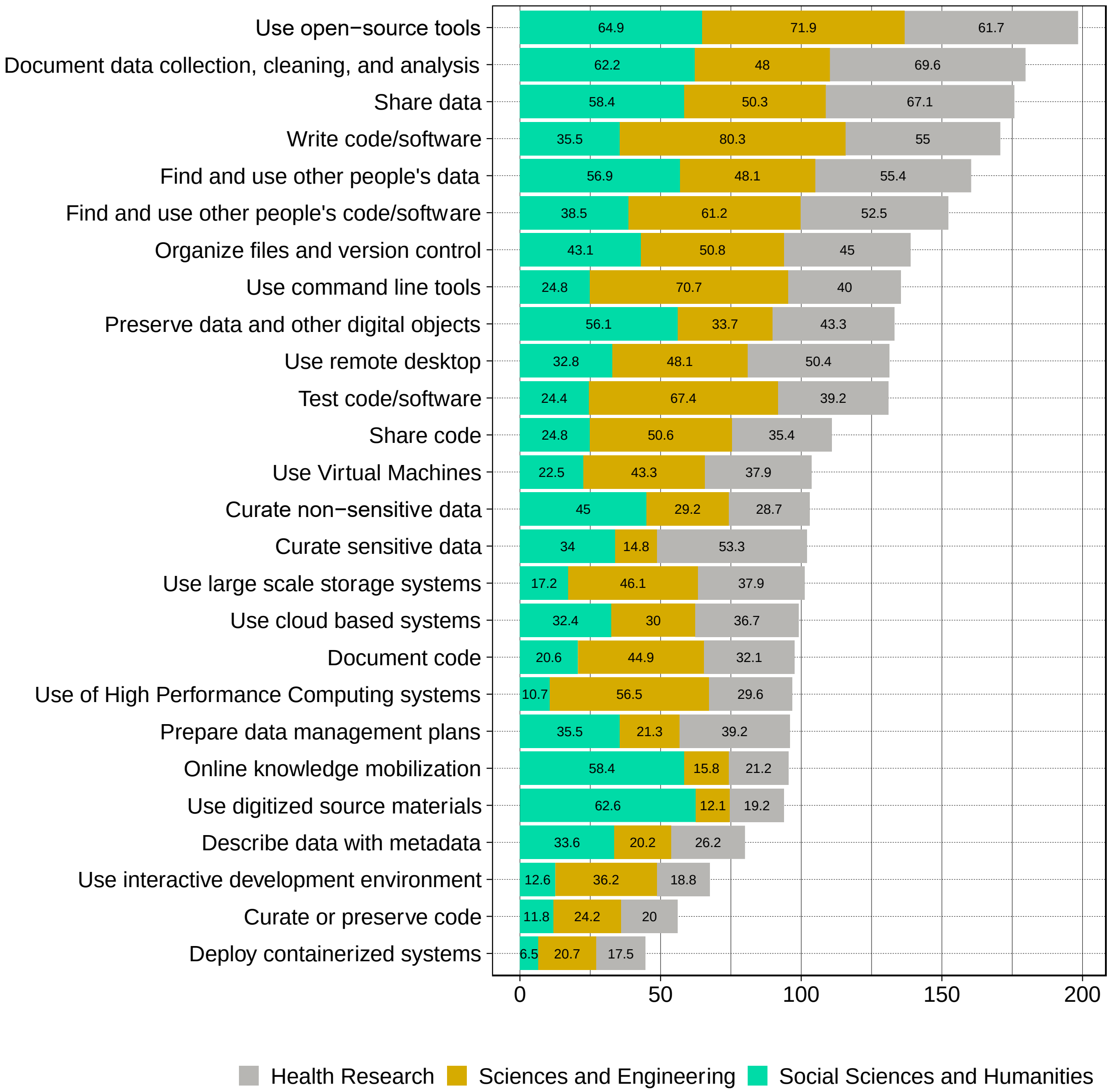

Figure 8. Survey result of one question in the Needs Assessment Survey: Which of these activities are part of your research workflows? Check all that apply. Multiple choice. (Social Sciences and Humanities, n=262; Health Research, n = 240; Science and Engineering, n = 644; Total = 1146).

In addition, some of the Position Papers that were submitted to the Alliance highlighted more focused RS-facing challenges that the Alliance could facilitate. The list below is a summary of key points from a selection of the Position papers: more detailed excerpts are included in Appendix B.

1. **General**

   a. Make it clear to researchers what (additional) services will be offered and make access to those services easier.
   b. Foster awareness, education, and culture change in an incremental, agile, non-intrusive manner.
   c. Provide a DRI framework which should include stable and reliable ARC infrastructures, easy access to tools to use national systems, provision of a central support team (including domain scientists, SW professionals, and HPC specialists) dedicated to: supporting Research Platforms; organizing calls and increasing funding to support the development and maintenance of research platforms; offering a long-term commitment to major Research Platform initiatives to bring them to the level of a sustainable national Service (research platform as a service, RPaaS).
2. **Capability:**
   a. **Funding**
      i. Establish a program to continue funding SW developers and RS/DRI support staff and RS teams at national, regional and local levels, including existing local research software development teams and expand scope to other universities across Canada.
      ii. Develop a program of secure, long-term and sustained investment in the development of new RS platforms, as well as the maintenance of RS.
      iii. Lower the barriers for access to DRI resources by funding research groups in system interoperability and data standardization.
      iv. Conserve researcher and institutional time by producing efficient systems for DRI funding administration and accountability.
      v. Develop a mechanism for fast turnaround resource application (e.g., RAC).
      vi. Develop a careful and thorough selection process along with criteria for choosing national RS or RS platforms, as well as a monitoring system to make sure they continue offering the expected service level.
   b. **Infrastructure**
      i. Large-scale domain-specific DRI that:
         1. provides a domain-specific view of resources (e.g., SW systems and tools, storage, computing, sharing, integrating and analyzing data, networks, etc.);
         2. provides 24/7 access, federated search portal, with uptime and system administration support;
         3. incorporates RS development, RDM and ARC;
         4. facilitates open research collaboration and scholarly communication (e.g., providing common annotation and provenance);
         5. enables the full spectrum of data;
         6. is agile and responsive;
         7. facilitates interoperability and standardization;
         8. offers a cloud-based environment;

9. provides respective building/modular blocks (e.g., updating and validating data types, user access management, registering projects, handling multiple data governance models, etc.);
10. enhances SW and data reuse;
11. establish a secured federated pan-Canadian DRI (ARC and RDM) system for accessing sensitive data, by complying with international sharing standards, such as GA4GH for genome data.

ii. Design effective interfaces between researchers and the system as a whole, represented by both people and machines.

iii. Provide a platform for Communities of Practice to serve infrastructure to specific groups, with EDI in mind.

iv. Provide a single, well-designed, usable, comprehensive, centralized administrative interface for managing user accounts, resource allocation, documentation, with links to as many administrative functions as possible, ensuring users are able to discover/access Alliance DRI resources.

v. Rationalize and improve interfaces for the system as a whole to reduce redundant activities (e.g., application, reporting) across the system.

vi. Explore the use of commercial cloud computing resources.

vii. Support multi-domain collaborative environments and tools to reduce information silos.

viii. Provide dedicated test and development platforms to test upgrades or on-board new experiments without disrupting the 24/7 production systems.

ix. Adapt and evolve large scale distributed computing and services, including cloud infrastructure and cloud-native technologies such as Kubernetes and Object Storage.

x. Support new analysis frameworks within the framework and ensure a consistent user experience

xi. Provide unified change management, notification and user support, e.g., by developing a uniform dashboard to collectively receive notification of system and other changes. Likewise, integrating performance and other monitoring across the federation sites.

c. **Research Data**

i. Serve as the research and data nexus among multi-stakeholders (government, academia, industry, indigenous nations, non-profit organizations, etc.) in the DRI ecosystem.

ii. Empower information systems by delivering data as a service over industry standard based APIs via service oriented architecture (SOA) and including comprehensive metadata.

d. Security

i. Extend the Alliance’s support interfaces to include Platform-as-a-Service (PaaS), particularly where it would improve efficiency, software quality, and security in the system as a whole.

3. Community

a. Events

i. Support the demand for RSE/RSes within local research support teams.
ii. Facilitate approaches to long-term recruitment and retention of HQP.
iii. Provide national disciplinary support for ongoing projects as well as new projects under preparation.
iv. Encourage and support the formation of community of practice (e.g., data/compute-intensive, Canadian RS community).

b. Training & Support

i. Provide national coordination of user support and training.
ii. Develop an RS strategy focused on training and adoption of best practices of RS (e.g., documentation) and promoting RS sharing and reuse.
iii. Facilitate local, regional, and national training support for trainees in SW development and in new technologies (e.g., ML, AI), as well as access to SW experts to support large-scale initiatives.
iv. Provide RS training that reflects the distinct context and practices in each specific domains.
v. Promotion of best practices for RS and DRI (e.g., FAIR, FAIR4RS, CARE).
vi. Ensure pan-Canadian and pan-disciplinary support interfaces are equitable regardless of region or specialization, breaking down trans-disciplinary silos.
vii. Foster scaffolding in under-served research areas through accessible platforms.
viii. Upskill IT professionals to manage the complex computing infrastructures, combined with domain-matter knowledge about research questions and research data.

4. **Coordination**

a. **Policy**

i. Create and promote incentives to share code.

b. **Governance**

i. Develop agile and adaptable strategies in a centralized fashion for ensuring accurate and informed communication tracking, together with timely responses.
ii. Develop a national strategy for coordination efforts, with peer agencies and/or service providers (e.g., CFI, Tri-Agencies, Genome Canada) in funding and resource allocation.
iii. Coordinate and collaborate with peer organizations and initiatives in funding and exploring new technologies.
iv. Enable customized communication flow between infrastructure team members and researchers for individual research project’s needs.
v. Strengthen institutional partnerships by connecting the campus-level DRI professionals, expertise and improvements to the national DRI.
vi. Build relationships with international entities and for-profit corporations.

vii. Devise a formal mechanism to receive suggestions, with continuous consultation.

viii. Provide iterative evaluation and revision, based on feedback from users and evaluation against the Alliance’s core values, to meet evolving needs

Appendix B: Position Paper Excerpts

1. **A Dynamic Environment for Medical Image Computing Research in Canada**[137]
    a. *Canadian researchers would benefit greatly if a system were in place that facilitated access to clinical medical images, from radiology, pathology, cardiology, neurology, etc. These images would be most valuable if they were accompanied with or linked to annotations and relevant clinical data providing details on diagnosis (including molecular biomarkers), treatment, demographic information and outcomes.*
    b. *Funding mechanisms that promote and encourage collaboration in the development of sharable code and tools*
2. **A Perspective on a Canadian Digital Research Infrastructure From: McGill Centre for Integrative Neuroscience McGill Centre for Integrative Neuroscience (MCIN)**[138]
    a. *Foster explicit, competitive career paths and retention for digital HQP working in universities.*
    b. *Sustain and grow existing, successful infrastructures and avoid "reinvention of the wheel".*
    c. *Prioritize interoperability and standardization for research infrastructure and supported outputs.*
    d. *Offer accessible and low-cost storage, processing and curation tools.*
3. **ACENET's Research Directorate White Paper for NDRIO**[139]
    a. *Undergraduate training bootcamps in basic programming skills, language switching (e.g., between C, Python, R, and Matlab), and "Software carpentry" should be available nationally both synchronously and asynchronously.*
4. **All Researchers Use Digital Resources: On Campus Support, Grants, Labs, and Equity**[140]
    a. *…actively support EDI by asking that proposed projects address potential biases in their design: the ableist, racist, sexist, and cisheteronormative tendencies of*

---

[137] Simpson et al., "A Dynamic Environment for Medical Image Computing Research in Canada", https://alliancecan.ca/sites/default/files/2022-03/ndrio_white_paper_medical_imaging_consortium.pdf.

[138] McGill Centre for Integrative Neuroscience McGill Centre for Integrative Neuroscience (MCIN), "A Perspective on a Canadian Digital Research Infrastructure From: McGill Centre for Integrative Neuroscience McGill Centre for Integrative Neuroscience (MCIN)", https://alliancecan.ca/sites/default/files/2022-03/ndrio.pdf.

[139] Rutenberg, "ACENET's Research Directorate White Paper for NDRIO Andrew Rutenberg on behalf of the ACENET Research Directorate", https://alliancecan.ca/sites/default/files/2022-03/acenet_rd_white_paper_for_ndrio.pdf.

[140] Estill, "All Researchers Use Digital Resources: On Campus Support, Grants, Labs, and Equity", https://alliancecan.ca/sites/default/files/2022-03/ndrio-white-paper-w-centernet-submitted_fixed-1.pdf.

many technologies and data models (including internet search algorithms, machine learning, AI, and hardware such as wearables)...

   b. *…ensure that students at all levels, postdoctoral fellows, and other non-faculty team members such as alternative-academics (alt-ac, such as lab managers and full- or part-time project staff) have access and are fairly credited and compensated for their time and labour.*

5. **Building a data science platform for better health**[141]
   a. *Data Architecture and Pipelines: Re-usable and extensible infrastructure that supports the data-to-analysis pipeline required for AI/ML and other computationally intensive analytics.*
   b. *Predictive Analytics Deployment: Software and developers to support the development and implementation of predictive analytics. Software components would enable the development, data collection and evaluation of the real-world use of the predictive analytics/AI and decision tools across geographies and institutions as well as for providers and patients.*
6. **Canada's Future DRI Ecosystem for Humanities & Social Sciences (HSS)**[142]
   a. *Easy, personalized web environment to deploy web applications, service should be offered with maintenance, and security updates (sysadmin services).*
   b. *Software solutions for long-term web hosting for scholarly communication*
   c. *Sandboxes and prototyping--particularly useful for graduate training*
   d. *An agile technological platform that responds to the variety of workflows/pipelines used by humanists in their work coupled with on-the-ground human support.*
7. **Canada's Path to a Global Open Research Commons**[143]
   a. *Going forward, data discovery services need to be reconfigured for notebooks and VREs so that searches among any data source can be conducted within the workspace.*
   b. *Achieving our vision for an ideal future for DRI in Canada is increasingly in reach because international coordination groups can help Canada identify and break down required tasks, while simultaneously developing a community of practice alongside the identified infrastructure and standards.*
8. **Considerations from researchers from Université Laval**[144]
   a. *Highly qualified personnel is the key element to support research software development.*
   b. *NDRIO should have a program to pursue the creation and support of research software teams to provide local support to researchers. NDRIO should also foster*

---

141 Rosella et al., "Building a data science platform for better health", https://alliancecan.ca/sites/default/files/2022-03/ndrio-white-paper-pophealth-final-rosella-dec14.pdf

142 Rockwell et al., "Canada's Future DRI Ecosystem for Humanities & Social Sciences (HSS)", https://alliancecan.ca/sites/default/files/2022-03/csdh_ndrio_whitepaper.pdf.

143 https://alliancecan.ca/en/canadas-path-global-open-research-commons.

144 Corbeil et al., "Considerations from researchers from Université Laval", https://alliancecan.ca/sites/default/files/2022-03/ul_researchers_ndrio_whitepaper.pdf.

a research software community that brings together institutional teams on the national level. This community will promote best practices in software development, encourage software reuse and help in bringing innovative software from teams at the scale of the national platform.

9. **Digital research infrastructure of two research laboratories in the field of structural dynamics**[145]
    a. *The promotion of best practices in RDM and RS truly implies a paradigm shift (1) to make all published articles fully accessible to the general public on central public repositories, (2) to promote the use of open-source software packages, (3) to disseminate in a sustainable manner open-source codes and research software, and (4) to ensure the continuity of research activities when a researcher or a student leaves a lab.*
10. **Digital research infrastructure to support federated computing on large scale biomedical datasets**[146]
    a. *To fully realize the benefits of digital health, we need national-scale data and software resources with privacy policies and hardware infrastructure to securely store, share and interpret medical data using AI and other advanced analytic tools.*
11. **Federating innovation in Digital Research Infrastructure accessibility**[147]
    a. *We propose a federated team with a global perspective on all fields of research that currently need access to our resources and that target needs specific to Canadian researchers for which solutions might not yet have been implemented internationally. Such a federated effort will require a team of full-time staff that can take ownership of promising initiatives before they get dropped by the local teams. The team's role would be to create and maintain a clear structure that would manage all levels of software accessibility, from supporting local initiatives to federal adoption and long-term support after the initial adoption phase.*
12. **Gaps in Digital Research Infrastructure for Canadian Digital Humanities Researchers**[148]
    a. *Provide support for the development of Research Software (RS) tools that facilitate the archiving and stewardship of DH data, and render them accessible online.*

---

[145] Batailly et al., "Digital research infrastructure of two research laboratories in the field of structural dynamics", https://alliancecan.ca/sites/default/files/2022-03/noirn_article_poly_mcgill4-1.pdf.

[146] Bourque et al., "Digital research infrastructure to support federated computing on large scale biomedical datasets", https://alliancecan.ca/sites/default/files/2022-03/digital-research-infrastructure_white-paper-submission_vf.pdf.

[147] Faure-Lacroix et al., "Federating innovation in Digital Research Infrastructure accessibility", https://alliancecan.ca/sites/default/files/2022-03/federating-innovation-in-dri-accessibility.pdf.

[148] Evalyn et al., "Gaps in Digital Research Infrastructure for Canadian Digital Humanities Researchers", https://alliancecan.ca/sites/default/files/2022-03/ndrio_dh_white_paper_digital_humanities_university_of_toronto.pdf.

13. **Interface Matters**[149]
    a. *Ensure the main NDRIO interface provides access to all of the functionality and information users need in order to access, manage, and report on their use of DRI.*
    b. *Ensure support interfaces are equitable regardless of region or specialization.*
    c. *Foster scaffolding in under-served research areas through accessible platforms.*
    d. *Provide a platform for Communities of Practice to serve infrastructure subgroups*
    e. *Extend NDRIO support interfaces to include Platform-as-a-Service, particularly where it would improve efficiency, software quality, and security in the system as a whole*.
14. **Large-Parallel Supercomputer Simulations – Frontiers in Canadian Research**[150]
    a. *…funding instruments are created that provide flexible support for creating, optimizing and adapting the specific software and big data ecosystems required by large-parallel simulations and their shared use by the broader community.*
15. **On Canada's Future Digital Research Infrastructure**[151]
    a. *Continue to grow the budget for both computer security personnel and for security training and education for researchers. (2) Enable more Platform-as-a-Service (PaaS) or Software-as-a-Service (Saas) offerings tailored to specific research needs. Such offerings could both improve research data security and reduce the burden on researchers of having to know about so many different threats, butwould require new investments in personnel to create and operate.*
    b. *The Research Software Engineer is recognized as a specialization in several other countries now (https://researchsoftware.org/). NDRIO should encourage the growth and recognition of this specialization within Canada, especially beyond Compute Canada or its successor organization. This is a specialization that should be given its due in research groups and universities.*
16. **On the Need for Local Research Software Development Funding**[152]
    a. *Having already established teams of highly skilled software developers at a number of universities across Canada … it would be an incredible setback to have these teams fall apart due to a lack of funding in the near future. …it is highly recommended that NDRIO establish a program to continue to fund local research software development teams at universities across Canada.*

[149] Antoniuk et al., "Interface Matters", https://alliancecan.ca/sites/default/files/2022-03/cwrc_interface_matters_ndrio_white_paper.pdf.

[150] Fernández et al., "Large-Parallel Supercomputer Simulations – Frontiers in Canadian Research", https://alliancecan.ca/sites/default/files/2022-03/ndrio_wp_on_niagara_scale_systems.pdf

[151] Dickson, "On Canada's Future Digital Research Infrastructure", https://alliancecan.ca/en/canadas-future-digital-research-infrastructure

[152] Schoenrock, "On the Need for Local Research Software Development Funding", https://alliancecan.ca/sites/default/files/2022-03/andrew-schoenrock-on-the-need-for-local-research-software-development-funding.pdf

17. **Persistent Identifiers in Canada – Position Paper**[153]
    a. *We recommend that NDRIO provides sustainable yearly funding for the Datacite Canada consortium and further promotes the adoption of DOIs for data sharing in Canada:*
        i. *Ensure integration of DOIs in all NDRIO-managed infrastructure*
        ii. *Require new research software projects funded by NDRIO to support DOIs*
    b. *We recommend that NDRIO provides sustainable yearly funding for the ORCID Canada consortium and further promotes the adoption of ORCIDs for data sharing in Canada:*
        i. *Ensure integration of ORCIDs in all NDRIO-managed infrastructure*
        ii. *Require new research software projects funded by NDRIO to support ORCID integrations*
        iii. *Encourage the adoption of ORCID iDs in grant application systems*
    c. *We recommend that NDRIO facilitate the adoption of, and support for, additional PIDs that help associate the researcher with other important elements of the research endeavor, such as Projects (RaID) and funding awards (Grant ID), thereby simplifying the discovery and impact of Canada's research outputs.*
    d. *We recommend that NDRIO encourages Canadian institutions to consistently use RORs when working within the PID frameworks (such as DOIs and ORCIDs) and supports the adoption of RORs in grant application systems.*
18. **McGill Response to NDRIO Call for White Papers on Canada's Future DRI Ecosystem**[154]
    a. *Canada needs a new national platform for digital infrastructure beyond high performance computing that also includes software resources, analytical tools and support for data management.*
    b. *Coordinate funding and resource allocation of digital research infrastructure with the major research funding agencies.*
    c. *Local support with sustained focus on training will expand access to the national digital research infrastructure system.*
    d. *Stable and predictable funding to host sites for operations of the national Advanced Research Computing platform.*
    e. *Adopt a Research Software strategy that incorporates additional focus on training and adoption of best practices as well as promoting a culture of software sharing and reuse.*
19. **Software support for ARC infrastructure**[155]

[153] Barsky, "Persistent Identifiers in Canada – Position Paper", https://alliancecan.ca/en/persistent-identifiers-canada-position-paper

[154] Crago, "McGill Response to NDRIO Call for White Papers on Canada's Future DRI Ecosystem", https://alliancecan.ca/sites/default/files/2022-03/mcgill_ndrio_driwhitepaper_20201214.pdf

[155] Taylor, "Software support for ARC infrastructure", https://alliancecan.ca/sites/default/files/2022-03/software-support-for-arc-infrastructure.pdf

a. *by cultivating in-house software support and development expertise on these open source software projects, the combined skill sets of system administrators and software developers can work together to identify a problem on a production system and resolve it in a future upstream version of the software, packaged and distributed as a standard release.*
b. *Contributing back to the FOSS (free and open source software) projects that we rely on for ARC infrastructure is a virtuous cycle that will make those projects better and healthier, benefiting their users and the research ecosystem, including us.*
c. *By funding software developer(s) who make code contributions back to the software projects that CCF ARC infrastructure relies on, we will foster productive working relationships with those open source communities, earn a say in the future development and implementation of features that may be of particular importance to us, allowing us to advocate for our priorities, and enabling us to conduct long-term planning of national ARC infrastructure deployments in conjunction with software development roadmaps.*

20. **Steps to Success in Ensuring DRI Engages and Mobilizes Humanities and Social Science Research**[156]
    a. *There is a need for infrastructure that is specifically designed for collaboration and software development, as well as the capacity to engage those with the expertise to facilitate this work, not just for the HSS Commons but for all Humanities, Social Science, and Information Science researchers and their DRI projects in Canada.*
21. **White Paper on Canada's Future DRI Ecosystem Subatomic Physics in Canada**[157]
    a. *…support the development of new research software and data management tools. No organization can predict the software and services that will be adopted by the researchers, and NDRIO should invest in many pilot projects rather than trying to select a system or service for everyone. Funds in these areas should be focused on our highest priority projects that have significant international impact based on peer review.*
    b. *NDRIO should explore the use of commercial cloud computing resources that could, for example, meet peak demands that exceed existing resources.*
22. **iReceptor – A case study in the challenges/opportunities in Canadian DRI**[158]
    a. *Consider the full spectrum of DRI funding and ensure that important funding streams that currently exist are not lost in the creation of NDRIO. Carefully*

---

[156] Siemens et al., "Steps to Success in Ensuring DRI Engages and Mobilizes Humanities and Social Science Research", https://alliancecan.ca/sites/default/files/2022-03/inke_ndrio-whitepaper_12-14-20.1.pdf

[157] Gezerlis et al., "White Paper on Canada's Future DRI Ecosystem Subatomic Physics in Canada", https://alliancecan.ca/sites/default/files/2022-03/sap-white-paper.pdf

[158] Breden et al., "iReceptor – A case study in the challenges/opportunities in Canadian DRI", https://alliancecan.ca/sites/default/files/2022-03/ireceptordriwhitepaper-final.pdf

consider continuity for the funding programs that currently exist and the platforms and projects they support.

b. *Build a DRI community within Canada. Focus service development on services that are likely to maximize the efficient use of the entire DRI ecosystem – help the researcher find the DRI pieces of the puzzle and assemble them into a solution that makes a difference.*

23. **Livre blanc sur l'infrastructure de recherche numérique : perspective de la Faculté des sciences de l'UQAM**[159]
    a. *Afin de stimuler la recherche, il est nécessaire de favoriser l'utilisation des logiciels développés par les chercheuses et chercheurs académiques. L'accès à ces outils de recherche serait facilité par une plateforme qui permet l'utilisation directe (en ligne) des logiciels. Le portail Canarie est un premier pas dans cette direction, mais en tant que répertoire il n'assure pas lui-même l'hébergement des logiciels. Il importe aussi de documenter le développement de ces logiciels, de la même manière qu'il faut documenter les autres aspects de la recherche. Certaines équipes utilisent actuellement Github pour garder une trace des modifications/ versions/ problèmes et solutions.*

24. **Digital Research Infrastructure for Canadian Astronomy**[160]
    a. *… NDRIO needs to invest in support staff. These staff members are essential to running the facilities, but can also contribute to the research projects. CANARIE offers programs to bring software developers into research groups, and NDRIO should offer similar programs.*
    b. *It is important to recognize the impact of international collaborations, and NDRIO should work to make connections with international counterparts.*

25. **Towards Research Platform as a Service (RPaaS) initiatives**[161]
    a. *In this document,we propose a framework to help NDRIO meet this objective,which should include a stable and reliable ARC underlying infrastructures, mid-level access tools to HPC systems, a central team dedicated to support Research Platforms,increase funding to support the development and maintenance of Research Platforms, and a long term commitment to major Research Platform initiatives to bring them to the level of an official Service (RPaaS).*
    b. *By taking over the Research Software mandate from CANARIE, it is very important that NDRIO continues this work in supporting the development of new and existing Research Platforms. There are certainly improvements to be made*

---

[159] Bergeron, "Livre blanc sur l'infrastructure de recherche numérique : perspective de la Faculté des sciences de l'UQAM", https://alliancecan.ca/sites/default/files/2022-03/uqam_fds_livreblancnoirn_vfinale.pdf

[160] Lovekin et al., "Digital Research Infrastructure for Canadian Astronomy", https://alliancecan.ca/en/document/259

[161] Gauthier et al., "Towards Research Platform as a Service (RPaaS) initiatives", https://alliancecan.ca/en/document/203

on this way forward, and we would like to take the opportunity to make some suggestions on how NDRIO could deliver on this goal:

i. *Organize calls to fund new Research Platforms led by researchers for development cycles of one to two years, followed by a maintenance period that may include further improvements in a similar way as CANARIE did. We believe that this model has proved to be productive and effective over the years.*
ii. *Organize calls to fund major improvements of Research Platforms already adopted by a community of users to expand their user base, again following similar calls by CANARIE.*
iii. *Allocate more money to fund these initiatives than what was previously available.This of course comes as no surprise, but it is clear that the need is there and we are convinced that there are many good quality projects just waiting to have the necessary funds to kick-off.*
iv. *Make sure the development teams include scientists the fields, software professionals and HPC specialists (most likely from existing local data centers) with a strong knowledge of the Canadian DRI. At the same time encouraging more synergy between the existing HPC staff in data centers and the research teams driving these initiatives.*
v. *Create a Research Software central support team to oversee the evolution of the various platforms with the objectives of promoting good practices, encouraging collaboration and software components reuse, and ultimately facilitating potential joint ventures between development teams.*

c. *Here we envision NDRIO playing an essential role in making RPaaS a reality for the Canadian researchers community, which will require a long term commitment for each supported RPaaS. So a careful and thorough selection process will need to be established along with a good monitoring system of the RPaaS to make sure they continue offering the expected service level. Here are some propositions on how NDRIO could support this transition towards RPaaS:*
   i. *Organize calls for 5 year long support among existing Research Platforms to acquire the status of RPaaS.*
   ii. *Criteria for choosing RPaaS will need to be carefully established.For Example: Number and category of active users; Usage of the Platform (sessions,jobs,storage,…); How well the Platform is appreciated by the community; Are there other tools/platforms doing the equivalent.*
   iii. *RPaaS would be evaluated every year, with the possibility of defunding the project through a mid-term assessment if things go wrong or if the Platform does not meet the RPaaS criteria anymore.*
   iv. *The RPaaS funding should only cover the cost to maintain the required level of services, including minor updates and small continuous improvements. Major Developments should go through the Research Platform development calls.*

v. *All RPaaS teams should include field specialists to directly assist the users.*

vi. *We think that 2-3 FTE per RPaaS for the maintenance would be a strict minimum to ensure staff redundancy and a decent quality of service (some PaaS could definitely need more). Some of the employees could be working part time on other components of the Canadian DRI, with dedicated time to the RPaaS projects.*

26. **HQP Pathways: Engaging the Canada's Different Disciplinary Models for HQP Training and Funding to Facilitate DRI Uptake in Canada[162]**
    a. *researcher-centric, service-oriented, and collaborative DRI maintenance and delivery that fits with disciplinary and funding norms of all researchers in Canada (in Humanities, Social Sciences, and Library Sciences this will mean providing systems administration and other support that cannot currently be covered by SSHRC or, in many cases CFI and provincial CFI matching programs).*
    b. *Out of the box provision of DRI and related support for common development and scholarly communication tools, such as a kubernetes cluster, GitLab instance, or CRM-backed website, rather than leaving individual research teams with the overhead and inefficiencies of installing and patching these systems themselves.*
27. **Submission to NDRIO's Call for White Papers on Canada's Future DRI Ecosystem from the Compute Canada Federation Subatomic Physics National Team[163]**
    a. *Support multi-Domain Collaborative Environments and Tools: … Some of the current tools used by CCF for internal collaboration (e.g., Slack, GitLab, staff wikis) have strict identity boundaries between those in CCF and those outside, preventing external collaborators from participating and creating information silos.*
    b. *Provide Dedicated Test and Development Platforms: As all the middleware and other supporting service components evolve continuously, there is a need for dedicated Test and Development systems within the CCF to be able to test upgrades or on-board new experiments without disrupting the 24/7 production systems.*
    c. *Support New Analysis Frameworks: New experiments are also exploiting new data analysis tools and techniques such as python-based Jupyter Hubs and projects coming out of the HEP Software Foundation . Adapting and evolving these within the CCF framework and ensuring a consistent user experience across the federated sites can be a challenge.*

[162] Siemens et al., "HQP Pathways: Engaging the Canada's Different Disciplinary Models for HQP Training and Funding to Facilitate DRI Uptake in Canada", https://alliancecan.ca/en/document/244

[163] Groer et al., "Submission to NDRIO's Call for White Papers on Canada's Future DRI Ecosystem from the Compute Canada Federation Subatomic Physics National Team", https://alliancecan.ca/en/document/211

d. *Provide Unified Change Management, Notification and User Support: the site-based dispersed change management, notification and support model is neither sufficient nor agile enough for SPNT projects that span multiple sites domestically or internationally. It is likely we need to develop a uniform dashboard for subatomic and similar projects to collectively receive notification of system and other changes. Likewise, integrating performance and other monitoring across the federation sites would be of benefit for the large projects.*

28. Gaps and Opportunities for NDRIO Support of Research Data Management[164]

a. *Element 3: Research software (RS), enabling researchers to access and use data*
   i. *There is a significant push by many organizations and entities to make data more accessible such as the GoC Open Government portal, Conservation Ontario, The Gordon Foundation Datastream and the FRDR. It would be helpful for NDRIO to provide additional leadership in this area to build an inventory of data portals available and harvested by FRDR, and additional metadata guidelines. … Additional tools for researchers include the provision of software such as MATLAB and geospatial tools. As well, application development that is funded through NDRIO should require an assessment of how the tool could be utilized by the broader research group. Often tools developed have a specific discipline focus which may limit the return on investment but sometimes it is necessary as some discipline do need domain specific tools.*

b. *Support Element 1: Highly qualified personnel, skilled people with the expertise to support the DRI system and help researchers make the most of cutting-edge tools*
   i. *It is clear that there is a great deal of expertise in the organizations that are now under the NDRIO umbrella. As described, there is a need for additional direct support of researchers tailored to their specific level of expertise. There is an opportunity to streamline services and make more efficient use of resources by taking advantage of the individuals within the system. … This resource should have an in-depth knowledge of the offerings of NDRIO and assist researchers in identifying appropriate resources to meet their needs.*

29. Digital Research Infrastructure in Astronomy[165]

a. *… To be useful, these portals require stable long-term funding on top of the base infrastructure funding. Without this stability, researchers will become reliant on systems that perpetually shift in their behaviour, and this will substantially impact productivity.*

---

[164] Dukacz et al., "Gaps and Opportunities for NDRIO Support of Research Data Management", https://alliancecan.ca/en/document/247

[165] Kavelaars, "Digital Research Infrastructure in Astronomy", https://alliancecan.ca/en/document/257

b. *The need for increased computational capacity is driven higher still by the development of new approaches to data analysis, in particular convolutional neural network (CNN)-based machine learning.*
c. *In a modern astronomical research community DRI should present to the community a domain-specific view of the resources, providing access to software systems and tools that are designed to meet the research goals of that community.*
d. *Canadian astronomy requires digital research infrastructure that can bring together these various digital pieces: computing, storage, networks, databases and software. Given the expressed science need for data collections to cut across sub-domains (such as X-ray, optical, optical, infrared and radio astronomy) a single astronomy domain aware science portal that enables use of the full spectrum of this data is needed. Such a system must be built in an agile way and respond on short timescales… Canadian astronomy must include a comprehensive science platform as part of our DRI.*

30. **"Good Things Come in Small Packets": How (Inter)national Digital Research Infrastructure can support "Small Data" Humanities and Cultural Heritage research**[166]
    a. *The vision promoted here is for a system that understands itself as part of an HCH research-to-publication workflow — in which data-infrastructure is a seamless part of a PID-based publication workflow.*
    b. *Streaming, API-access to data stored in DDRs — the precise mix is less important than the understanding that a DRI should support the publication of HCH research data in context.*
    c. *What is crucial, however, is training and community work to make the change from current practices to an Open, FAIR (and, as appropriate, CARE)-complaint DRI ecosystem.*
    d. *Rather than proposing specific tools, we'd argue that NDRIO should consider an environment scan — looking at specific use-cases with discipline experts and comparing how such use-cases are or are not supported by existing DRIs. It is striking the degree to which most of the tools, services, and resources already exist, but are distributed unevenly across different systems.*
    e. *As a national agency in a country where there a) appears to be reasonably good integration and collaboration among the three main agencies; b) is an appetite for interdisciplinary work (e.g., Frontiers); and c) funding exists for the whole research cycle from initial idea (the various discovery and development grants) to publication (ASJ), we believe that NDRIO is well-situated internationally to take the interdisciplinary approach proposed here.*

[166] O'Donnell, "Good Things Come in Small Packets": How (Inter)national Digital Research Infrastructure can support "Small Data" Humanities and Cultural Heritage research", https://alliancecan.ca/en/document/195

31. **A Standards-Based Digital Infrastructure for Secure Sharing of Human Biomedical Research Data**[167]
    a. *… break down the barriers for responsibly sharing human biomedical research data by building a national digital infrastructure for storing, sharing, integrating and analyzing large biomedical research data sets under a responsive and responsible data governance regime that balances human subject's privacy needs against the imperative to make collected data as widely available to the scientific research and clinical communities as possible.*
    b. *We propose the creation of an open source software suite for submitting, storing, finding, integrating, and responsibly sharing the fruits of Canadian biomedical research .*
32. **Empowering Information Systems and Fostering Metadata Driven Data Management**[168]
    a. *Modernizing and strengthening infrastructure by empowering information systems is however fundamental and supports progress across the board. This requires software to have access to the same knowledge as their human counterparts through service oriented architectures and comprehensive metadata.*
33. **Standardized, linkable, analysis-ready environmental data for understanding and preventing disease and building resilience to climate change**[169]
    a. *To bridge the gaps revealed through CANUE's experience to-date, key infrastructure (software and system) capabilities include:*
        i. *Automated creation of digital metadata providing structured variable dictionaries and assigning unique and persistent identifiers for each dataset on upload to the portal.*
        ii. *Interfaces for conducting complex spatial and metadata searches to facilitate data exploration, interpretation, and identification of variables/datasets of interest.*
        iii. *Software to process new datasets not yet included in the data portal via scheduling functions that automatically and regularly pulls source environmental data available on open platforms.*
    b. *Digital infrastructure is clearly critical to supporting research needs; however, expertise to support the infrastructure cannot be ignored. While computer scientists, data scientists and data managers are all essential, the resulting data are best improved and shared when those working on the infrastructure also have content expertise across disciplines. Researchers accessing the data can*

[167] Stein et al., "A Standards-Based Digital Infrastructure for Secure Sharing of Human Biomedical Research Data", https://alliancecan.ca/en/document/286

[168] Heus, "Empowering Information Systems and Fostering Metadata Driven Data Management", https://alliancecan.ca/en/document/252

[169] Brook et al., "Standardized, linkable, analysis-ready environmental data for understanding and preventing disease and building resilience to climate change", https://alliancecan.ca/en/document/217

bring some of this, but all are better-served when those working day-to-day building the infrastructure have a deep understanding of the substantive content and the research questions it/they are supporting. Developing and motivating this expertise can thus not be over-looked.

Appendix C: Current and Future State

The RSoF model was applied by considering the topics and sub-topics from the perspective of the Alliance (defined as the Alliance services and resources, including those provided by the Federation and Network of Experts), rather than via the lens of the research software community at large. Taking this approach can be challenging, and it can be difficult to reconcile a status for the Alliance vs the broader community, especially when the community may be further along than the Alliance. For example, the Alliance is currently not far along in promoting the use of standards in the development of research software, but specific research software platforms may well be very advanced in this context. The goal is to consider what the Alliance has done, and can do going forward, to support the community in each aspect of the model. It is also important to bear in mind that the RSoF is intended to guide the discussion and development of support programs in the 2025-30 period, and is not a rigorous review and ranking of every aspect of the research software ecosystem.

The aspects of the research software ecosystem are classified by the research lifecycle stages (Envision, Plan, Develop/Acquire, Use, Share/Reuse/Preserve/Discard), and within those categories by Topic and Sub-Topic. These categories were derived from version 1.5 of the NIST's RDaF matrix, and modified to reflect research software: in some cases this meant adding elements, or dropping some, as the goal was for the RSoF to be complementary to the RDaF, and not overlapping. The RSoF model considers 50 topics (presented in Table 2) and over 350 sub-topics. With this as the starting point, the process was:

1. Each Sub-topic was classified as In-Scope (relevant for the 2025-30 mandate), Not In Scope (relevant for the 2025-30 mandate), or Future (relevant for a post 2025-30 mandate).
2. Each Sub-topic was ranked on a scale of 0-5 with respect to the **current state** maturity level (according to the CMMI below in Figure 9), as well as an indication of the reason for the maturity designation.
3. Each Sub-topcic was ranked on a scale of 0-5 with respect to the **desired future state** maturity level (according to the CMMI below), as well as an indication of the reason for the maturity designation, as well as a description of the desired future state.
4. Each high-level Topic was then assigned a **priority** (0-5), which reflects an estimate of how important that Topic and the associated Sub-topics would be in the 2025-30 mandate.
5. After the rankings were applied, the type of potential support (using the categories from the Possible Approaches section of this document) was used to apply one example of an approach to achieving the desired state for that Sub-topic. It is important to note that there may often be more than one appropriate form of support for a specific aspect: the goal with this was to suggest one possible approach as an example.

6. The matrix includes additional fields which may be used by the WG, or the Alliance's internal team to provide additional information for the final 2025-30 submission to ISED. These include: Model, an indication of the Sub-Topic reflecting a National, Regional, Institutional, or International focus; Funding Level, an indication of the level of funding or support (< $250K, < $1M, < $5M, > $5M) needed to move to a desired state; Mandate Year Start, a potential start for an initiative from Years 1-5; Mandate Years, the potential length in years of an associated initiative from 1-5 years.

Figure 9. The CMMI Maturity Model

The table of 50 topics from the RSoF are presented below, grouped by RS Lifecycle stage, listing the label for the topic, the current and desired future state, as well as the topics assessed priority. A series of charts that illustrate the effort needed to move from the current state to the desired future state are also provided here. The full RSoF model has additional information not included here, including rationale for the CMMI ranking, detailed topic descriptions, etc. The full suite of Sub-topics in the RSoF was considered in combination with our assessment of international best practices (see Research Software in Context) and the RS Landscape in Canada to arrive at a high-level set of Recommendations for how the Alliance can best support the RS community in the 2025-30 mandate. Those Topics which had a priority of 5 or 4 (Highlighted in Table 2) were used to focus the Recommendations for 2025-30.

Table 2. RSoF Topics with their current and desired future state.???

Lifecycle Stage	Label	Topic	Current State	Future State	Priority
Envision	Governance: Strategic	RS Governance: Strategic & Qualitative	.8	3.2	5

	Governance: Legal	RS Governance: Legal and Regulatory Compliance	1	2.9	5
	Culture	RS Culture and Reward Structure	.4	2.2	5
	Education	Education and Workforce Development	.6	2.5	4
	Sustainability	Resources—Allocation and Sustainability	.7	4.0	5
	Community	Community Engagement	.3	2.8	5
Plan	Control	Chain of Control (Custody of Software)	1.5	2.8	2
	Financial	Financial Aspects of Planning	0.1	4.0	5
	Planning	RS Planning	1.3	4.0	4
	Training	Training support for SW	1.5	3.0	5
	FAIR	FAIR	1.8	3.8	3
	Metadata	SW Metadata Considerations	1.2	3.0	3
	Architecture	SW Architecture	1.5	3.5	2
	Standards	Standards	0.3	2.3	4
	Assessment	Assessment	0.5	4.0	5
	Outreach	Communication and Outreach	0.6	3.0	3
RS Develop	Layer	SW layer	1.5	3.0	1
	License	SW license	0.0	3.0	1
	Distribution	SW distribution	1.5	4.5	1
	Purpose	SW purpose	0.1	2.4	2
	Dev Approach	SW development	2.0	3.5	4
	Dev Process	SW development processes	0.2	4.0	5
	FAIR	FAIR Principles	0.0	3.0	5
	Standards	Community based standards	0.0	3.7	5
RS Use	Types of RS	Types of RS	0.0	2.5	4
	Discovery	SW discovery	0.2	4.1	4
	Management	SW Management	1.7	4.0	5

	Prep	Preparation and Pre-Processing Methods	0.0	1.0	1
	Analysis	Analysis Methods	0.0	1.0	1
	Modeling	Modeling	1.3	3.3	2
	Workflows	Workflows and Middleware	0.8	3.1	4
	Provenance	SW Provenance	0.8	2.2	3
	Licensing	SW Licensing	3.0	4.0	3
	Execution	SW execution environment	1.3	3.8	1
Share Reuse	Publishing	Publishing	0.7	2.5	5
	Dissemination	Modes of Dissemination	0.4	2.6	5
	Attribution	Attribution	0.4	3.1	5
	Sharing	Modes of Sharing	2.0	3.8	4
	Access	Access	0.0	1.7	3
	Legal	Legal and Licenses	0.0	2.3	3
	Protection	Levels of Protection	0.8	2.8	4
	Architecture	Architectures for Application, Use, and Reuse	0.0	2.5	2
Preserve/Discard	Preservation	Criteria for Preservation	0.0	2.3	4
	Sustainability	Sustainability	0.7	3.0	5
	Storage	Storage and Preservation	0.8	2.8	3
	Mobility	Moving Data from One Service to Another across Organizations	0.0	2.3	2
	Retention	Retention and Disposition Schedules	0.0	2.0	1

The diagrams below (Figure 10 to 15) present the current and future state for each of the stages of the RS Lifecycle, split into relevant key topics from RSoF. Current state is represented by the blue line while the desired future state is represented by the red line. The effort to move a topic from current state to future state is represented by the difference along the radial axis between the blue and red lines.

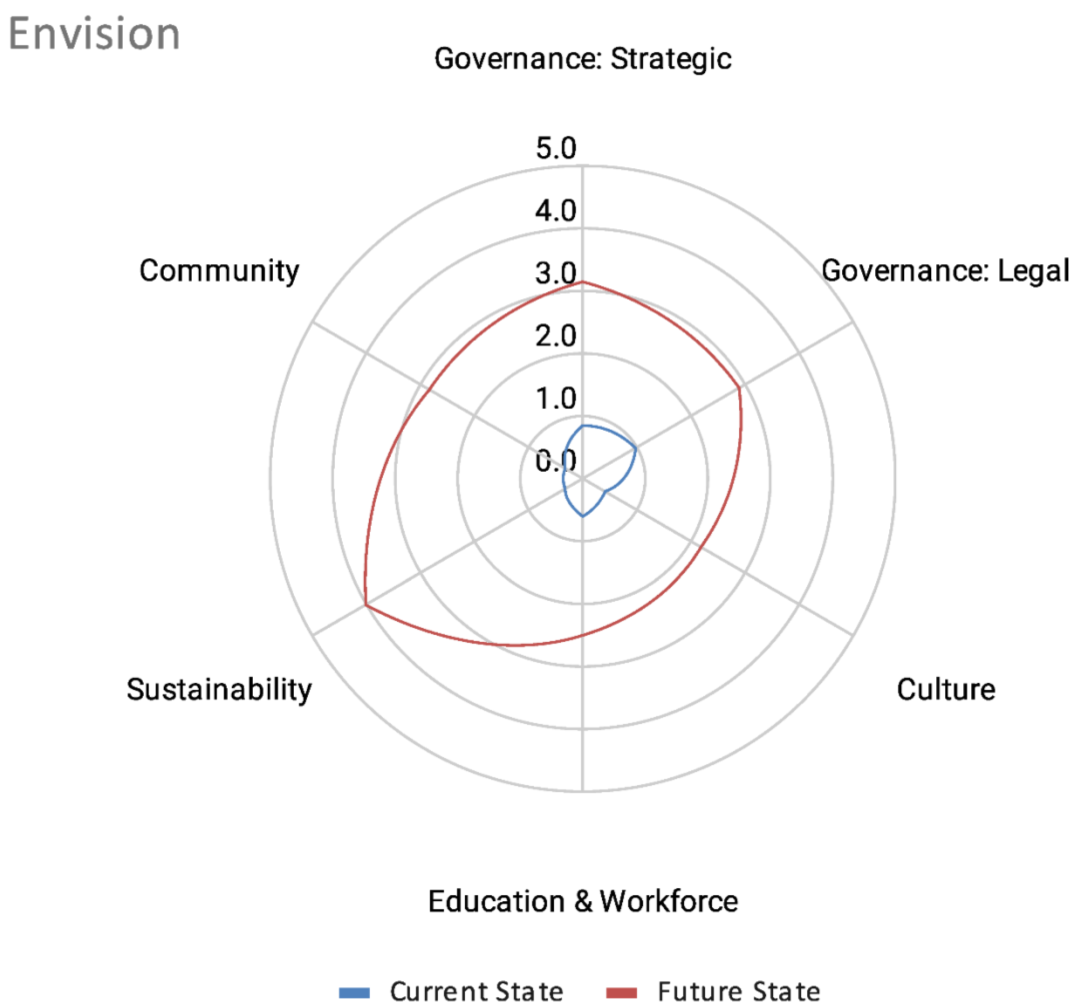

Figure 10. High-level summary of RSoF Topics for the Envision stage of the Research Software Lifecycle.

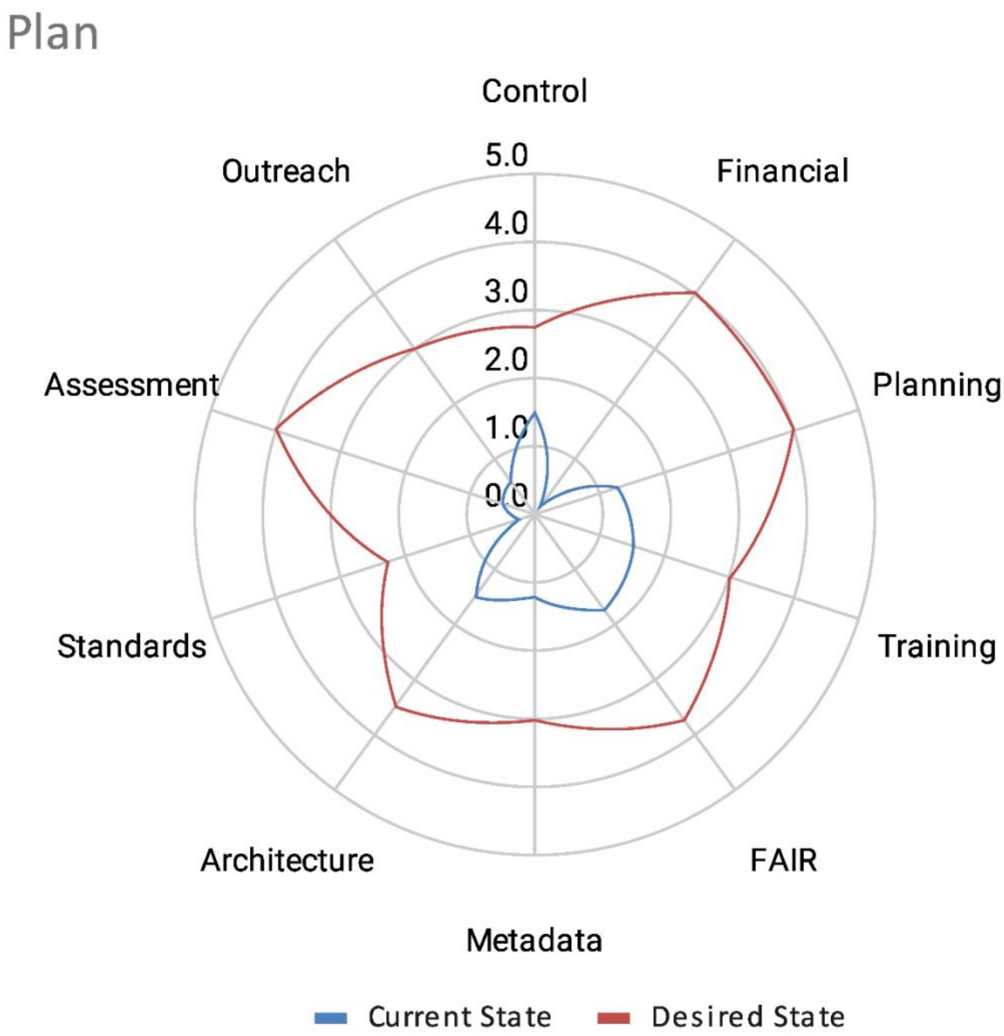

Figure 11: High-level summary of RSoF Topics for the Planning stage of the Research Software Lifecycle.

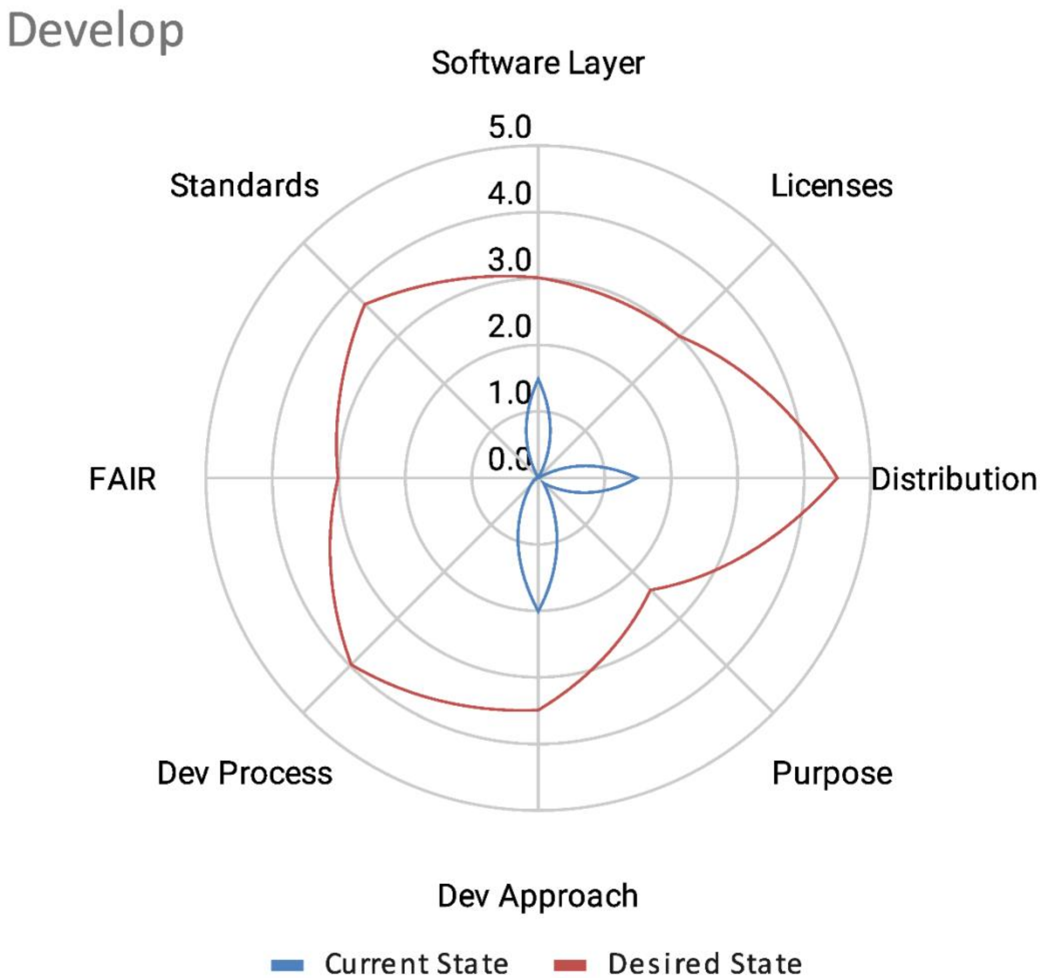

Figure 12. High-level summary of RSoF Topics for the Development stage of the Research Software Lifecycle.

Use

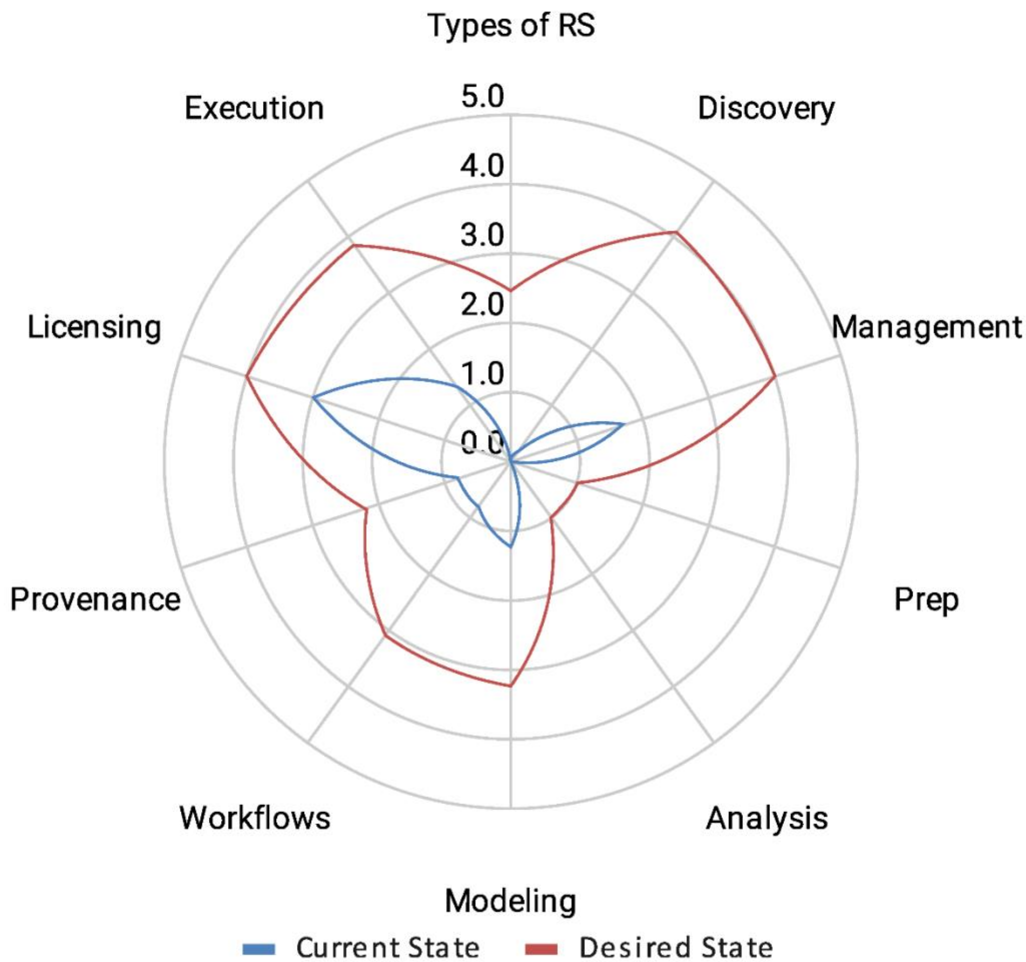

Figure 13. High-level summary of RSoF Topics for the Use stage of the Research Software Lifecycle.

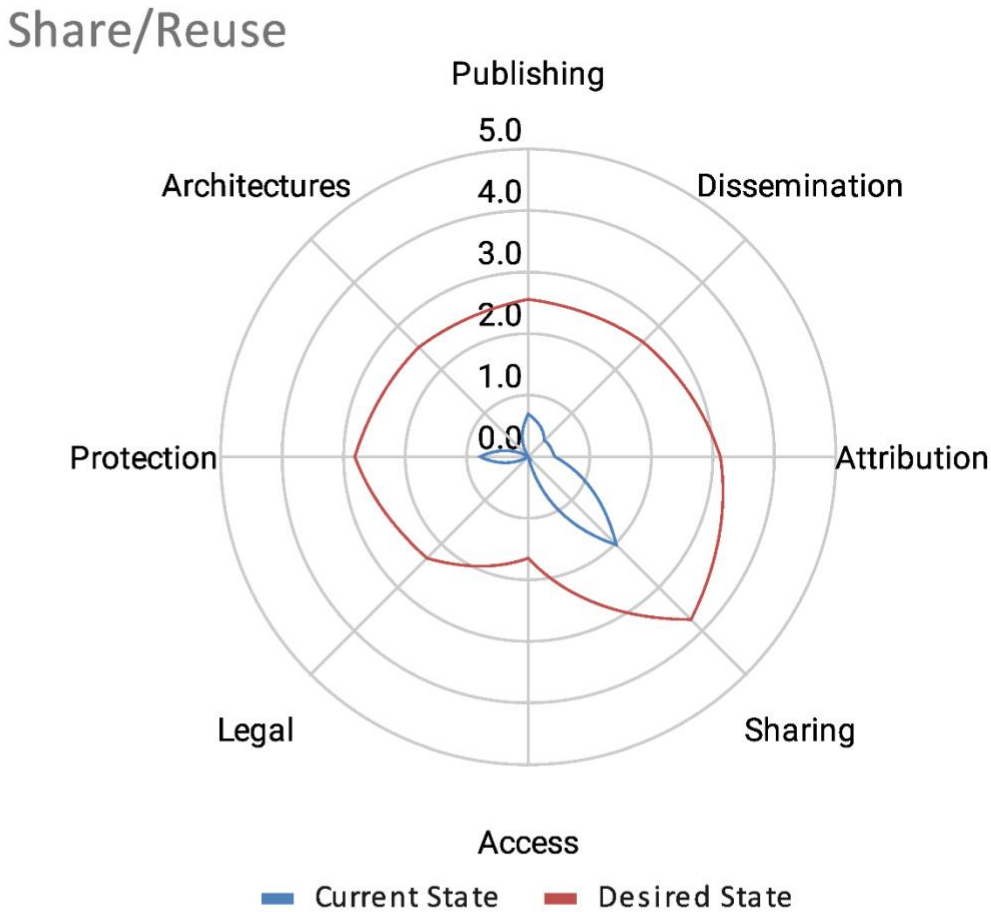

Figure 14. High-level summary of RSoF Topics for the Share/Reuse stage of the Research Software Lifecycle.

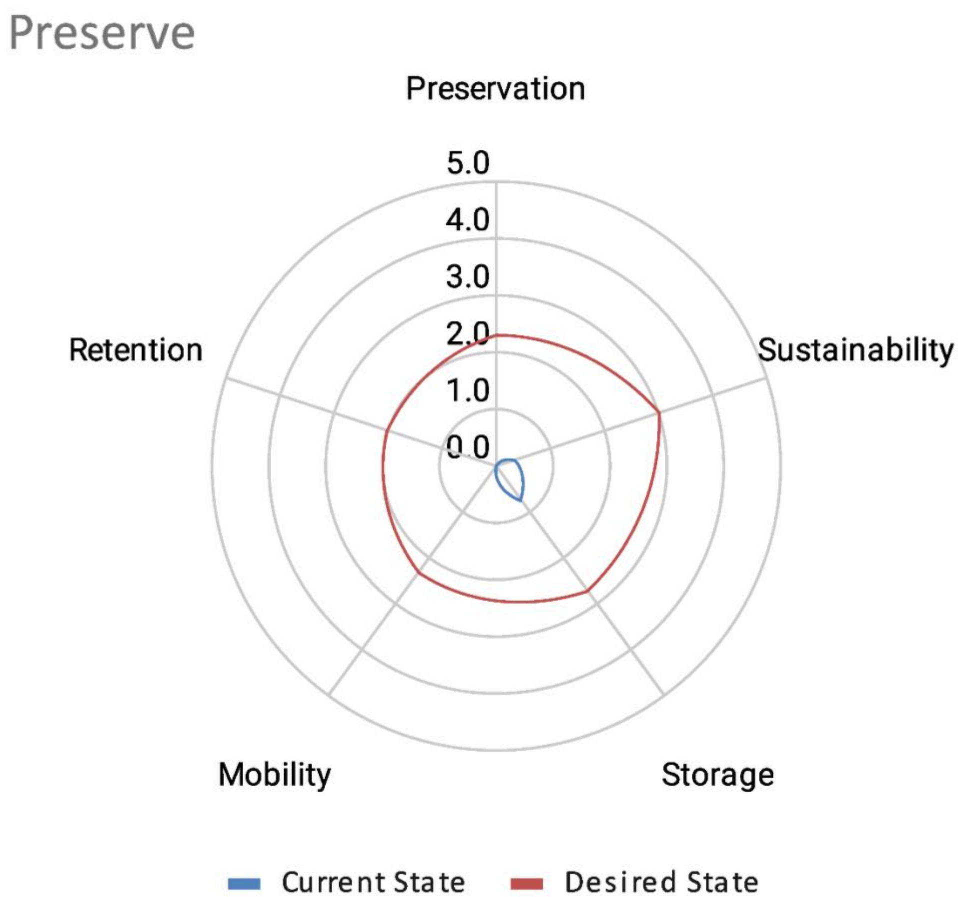

Figure 15. High-level summary of RSoF Topics for the Preserve stage of the Research Software Lifecycle.

Appendix D: Metrics for the Assessment of RS Platforms

In this section we summarize a number of initiatives to establish metrics that capture the success of RS Platforms.

CANARIE Metrics

CANARIE funded projects were required to provide a set of common metrics which CANARIE used to assess the success of the RS Platforms that were developed through their programs and to report on those successes to their stakeholders. CANARIE projects were required to report on the following:

- Number of Canadian Research Teams
- Number of International Research Teams
- Number of funder/publisher/aggregator recommendation for platform
- Number of platforms/repositories integrated
- Number of demos/presentations/showcases/paper/posters about project
- Uptime of service
- Accesses to the service

NIH Metrics

In 2021, the National Institutes of Health (NIH) Data Resources Lifecycle and Metrics Working Group and Metrics for Repositories (MetRe) subgroup released "Metrics for Data Repositories and Knowledgebases: A Working Group Report"[170] as an exploration of the current landscape of biomedical data repository metrics. A summary of the metrics discussed in this document are given below:

- User behaviour characteristics
    - Number of users
    - Page views
    - Downloads
    - Location (IP address)

[170] Metrics for Data Repositories and Knowledgebases: Working Group Report, https://datascience.nih.gov/news/metrics-for-data-repositories-and-knowledgebases-working-group-report

    - New vs Returning (for defined period e.g., 3 months)
    - Dataset submitters
    - Visit frequency (daily, monthly)
    - Data access requests (for defined period)
- Scientific contribution
    - Number of Projects/Studies
    - Number of Cases/Subjects
    - Total publications
- Operations
    - Storage costs
    - Cost/dataset (storage)
    - Total Hardware coss
    - Total Download costs

Elixir Metrics

Similarly, Elixir has established a “Plan For Collation Of Metrics And Quality Data At The Elixir Hub”[171] for Elixir Core Data resources. The report states “The collection of indicators from the Core Data Resources allows ELIXIR to demonstrate the need of sustained infrastructure funding to funders and stakeholders and monitor progress, trends and usage over time.”

A summary of the metrics recommended by Elixir are:

- Scientific focus/quality
    - Archive VS Knowledge base (check box)
    - Scope statement
    - International dimension description
    - Staff effort
        - Curators
        - Bioinformaticians
        - Technical staff
- Community
    - Overall usage
        - Visits/sessions (yearly)
        - Page view (yearly)
        - Unique users/IP address (yearly)

[171] Stockinger et al., “Plan for collation of metrics and quality data at the ELIXIR Hub”, Zenodo, March 8, 2018, https://doi.org/10.5281/zenodo.1194122

        - Hits/requests (yearly)
        - Downloads (yearly)
    - Usage in literature
        - Number of citations of name of resource mentioned in Europe PMC.
        - Number of citations of Accessions Numbers mentioned in Europe PMC.
        - Number of citations of formal paper describing the resource
        - Dependency of other resources
    - Quality of service
        - Identifier use (PIDs???) (free text)
        - Data throughput
            - Data entries
            - Data size
        - Technical performance
            - Uptime
            - Response time
        - Use of standards (free text)
        - Links to provenance
        - Data availability
        - Customer service
    - Legal and funding, governance
        - Scientific advisory board
        - Open Science
        - Privacy Policy
        - Ethics Policy
        - Sustainable support and funding
    - Impact and translational stories
        - Counterfactual statement (what happens if it isn't around)
        - Accelerating science
        - Translational data

Appendix E: Review of National and International RS Funding Programs

ReSA defines two high level categories of research software:[172]

1. **Category A**: Within a research project as a by-product to do the research
    - Most funding for research software fits in this category; however, it is difficult to quantify the amount of investment. The funding quantum is certainly significant, with research showing that ~20% of National Science Foundation (NSF) projects (totalling USD$10b) over 11 years discussed software in their abstracts (Katz 2021), software-intensive projects are a majority of current publications (Nangia and Katz 2017), and 33% of research produces new code (Bello and Galindo-Rueda 2020, fig 3.4). General research funding does not usually name research software as a potential output, although there is increasing inclusion of research data, which include mention of tools for data analysis (which can include research software).
2. **Category B**: Through intentional development of a software product for general use in research by one or more projects.
    - Funding for research software through this mechanism is provided by some funders. However, it can be difficult to identify these as opportunities for research software as the funding is often provided within broader frameworks, such as digital infrastructure, technology innovation, and open science.

It is these later programs (Category B) that we discuss below.

The list below provides examples of RS funding programs from the national and international community, and provide exemplars and examples of approaches that can help the Alliance address the desire to move to the future state. These examples are listed in the Recommendations section in the context where they could provide models for the Alliance's own programs.

[172] Barker et al., "Overview of Research Software Funding Landscape". Zenodo, February 24, 2022, https://www.doi.org/10.5281/zenodo.6102487.

Capability

New Components / Platforms

For each type of funding approach listed below we describe what it means, and provide examples from existing international funders. Where previous CANARIE funding approaches match current or emerging examples, we also list those to illustrate the excellence of some previous RS funding programs in Canada.

Innovation Calls

1. eScience Center Open eScience Call 2023 (OEC 2023)
    a. Support state-of-the-art and innovative research that requires the development and application of advanced research software. (Netherlands)
2. SAGE Publishing: Sage Concept Grants
    a. This program provides funding for innovative software solutions that support research in the social sciences. The grants are designed to fund new technological solutions that support the adoption, development and application of established and emerging research methods, including quantitative, qualitative, mixed, and computational methods. This could include:
        i. Supporting social scientists to collect and work with new types of data.
        ii. Supporting social scientists to use and apply new methods.
        iii. Reducing the time researchers spend on the collection, cleaning, or filtering of data, or any other part of the research process that is not publishable.
        iv. Enabling social scientists who do not have programming skills to analyze data at any scale.
        v. Enabling researchers to get started on their research without much training.
        vi. Improving existing social science methods or tools.

    One grant of £15,000 will be awarded to scale up a prototype, and a number of £2,000 grants will support concept testing for early stage ideas. (Global)
3. CFI Cyberinfrastructure Challenge 1
    a. Challenge 1 of the Cyberinfrastructure Initiative involved two competitions through which we invested in research data infrastructure projects that enabled 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 to enhance the community’s capacity to conduct leading-edge research.

Existing Platforms

Development

1. Enhancement
    a. DFG Increase the Usability of Existing Research Software
        i. This interdisciplinary call for proposals pursues three sub-goals which together seek to raise the maturity level of research software so that it can also be used by researchers other than those developing it, and to simplify its further development: usability and impact, quality assurance, and further development. (Germany)
    b. CANARIE Onboarding New Teams Call
        i. CANARIE announced today the selection of 13 successful projects from its latest Research Software funding call. This funding will enable research teams to adapt their existing research platforms for re-use by other research teams, including those working in different disciplines. As a result, new research teams from across Canada will be able to re-use previously funded and developed software to accelerate their discoveries.
2. Sustaining
    a. eScience Centre Call for Sustainable Software 2023 (SS 2023)
        i. Enhance successful research software. Supports communities of researchers who require their software to meet higher quality standards to ensure the continuity and advancement of their research in the longer term. Eligibility: researcher employed at a Dutch organization, hold a PhD, hold a permanent contract, commitment to the project work for half a day per week on average for the duration of the project. (Netherlands)
    b. Department of Energy (DoE) Seed Collaborations for Software Sustainability
        i. The Advanced Scientific Computing Research (ASCR) program seeks to seed collaborations focused on software sustainability. Applications should articulate a long-term vision for sustaining a substantial part of the ecosystem of software for scientific and high-performance computing and should describe how their efforts might be rapidly scaled up in the future. (USA)
    c. Global Biodata Coalition Calls 1 and 2
        i. The Global Biodata Coalition has had two calls to identify “global core biodata resources”, with the future goal of defining more effective approaches to funding their enhancement and sustainability. This effort has been one of the most important developments in the establishment of a framework for defining core platforms and resources, and determining how to fund them going forward.

National Integration / Interoperability

1. CANARIE RDM Programs 1 and 2

a. Prior to the move of CANARIE RS and RDM programs to the Alliance, there were two funding calls which had a focus on the development of RDM software and platforms that facilitated the adoption of internationally recognized best practices, as well as global interoperability. The platforms that were successful in these calls included some of Canada's leading RS platforms and tools. The 2nd Call had the primary intention to "Extend the Interoperability of Canadian Research Data with National and Global Systems".

Thematic Approach

Domain-Specific

1. CZI Single Cell Biology Data Insights
    a. The Chan Zuckerberg Initiative invites applications for the second of three cycles for 18-month projects focused on advancing tools and resources that make it possible to gain greater insights into health and disease from single-cell biology datasets. (Global)
2. CIHR/EU H2020 - International flagship collaboration with Canada for human data storage, integration and sharing to enable personalised medicine approaches
    a. To build a collaboration of stakeholders in Europe and Canada in the domain of repositories storing and sharing human –omics data that will create a framework for long-term cooperation. In order to do so, this programme aims to enhance and standardise data deposition, curation and exchange procedures thus ensuring better data reuse and increased benefit to the scientific communities worldwide. The selected projects should build on the data quality metrics, standards and access policies developed by major international initiatives (e.g., IHEC, ICGC[[https://dcc.icgc.org/]], IHMS[[http://www.human-microbiome.org/]], MME[[http://www.matchmakerexchange.org]])

Challenge Focused

1. Open Scholarship
    a. NASA High Priority Open-Source Science (HPOSS)
        i. Program goals include making science more accessible, inclusive, and reproducible. Proposals must be for new work to develop technology that will support open-source science, such as development of data formats, software, frameworks, or libraries. Awards of $100,000 are available for one year. (USA)
    b. CHIST-ERA Open & Re-usable Research Data & Software
        i. This call tackles the challenge of open research data and software from the perspective of their possible reuse. The objective is to create the conditions for research in any domain based on open or shared data and software. (the project consortia must have a minimum of 3 partners from at least 3 of the following countries: Belgium, Brazil, Czech Republic,

France, Latvia, Lithuania, Luxembourg, Poland, Romania, Slovakia, Switzerland, Turkey, United Kingdom)

2. Open Source
    a. CZI Essential Open Source Software
        i. The Chan Zuckerberg Initiative invites applications in support of open source software projects that are essential to biomedical research. The goal of the program is to support software maintenance, growth, development, and community engagement for these critical tools. (Global)
3. Data Interoperability
    a. CANARIE RDM
    b. CANARIE, a vital component of Canada's digital infrastructure supporting research, education and innovation, today announced nine successful recipients of its Research Data Management (RDM) funding call, announced in May 2018. This new funding will enable research teams to develop software components and tools to enable Canadian researchers to adopt best practices in managing data resulting from scientific research. Data management practices impact the entire research lifecycle, from project planning and execution, to backing up data as it is created and used, and finally to its long-term preservation after the investigation is complete. RDM best practices help ensure the protection of data during the research lifecycle and beyond, and help meet the increasingly stringent requirements of research ethics and reproducibility.

Multi-National Funding Calls

1. CIHR/EU H2020 - International flagship collaboration with Canada for human data storage, integration and sharing to enable personalised medicine approaches
    a. To build a collaboration of stakeholders in Europe and Canada in the domain of repositories storing and sharing human –omics data that will create a framework for long-term cooperation. In order to do so, this programme aims to enhance and standardise data deposition, curation and exchange procedures thus ensuring better data reuse and increased benefit to the scientific communities worldwide. The selected projects should build on the data quality metrics, standards and access policies developed by major international initiatives (e.g., IHEC, ICGC[[https://dcc.icgc.org/]], IHMC[[http://www.human-microbiome.org/]], MME[[http://www.matchmakerexchange.org]]
    b. Example Canadian platforms: iReceptor, IRIDIA, …,

AI/DL/ML-Facilitated Research

1. DOE FAIR Data to Advance Artificial Intelligence for Science
    a. A funding call in 2020 aimed at "making artificial intelligence (AI) models and data more accessible and reusable to accelerate AI research and development (R&D)". The funded projects covered a range of topics including high

performance computing, materials science, high energy physics, and microbial science, but all had elements of RS development.

Community

Events

1. Conferences
    a. CANARIE Research Software Workshop
        i. An annual event (ceased in 2022) which was a gathering of research software teams funded under the CANARIE programs. The event was enhanced with software developers from CANARIE's RDM calls. The 2-day event was considered a success by all participants, and until the RS pillar moved to the Alliance, it was the only focused gathering of RS teams in Canada.
    b. RSE Conference
        i. Starting in 2016, RSE (Research Software Engineers) Conference is a turning point in the RSE movement, which not only brings RSEs together but sparks the creation of related movements in countries and around the world (Germany, Netherlands, Nordic, Australia and New Zealand, US, etc.). The conference has continued to grow in popularity every year. This year, the RSE Conference 2023 (RSECon2023) will be held at Swansea University 5-7th September, 2023.
2. Alliance Hosted
    a. DRI Connect Event
        i. The DRI Connect event replaced the TECC event, which was hosted (typically twice a year) by Compute Canada, and was designed to bring together individuals who were part of the broader Compute Canada community. While not RS-focused, it did include specific sessions with a RS theme.
3. Workshops / Hackfest
    a. ReSA Funders Forum Workshops
        i. The 1st Workshop was held in November 2022, and the 2nd will be held in Montreal in September 2023, and hosted by the Alliance. This event brings together international RS funders to discuss how to improve funding for RS globally.
    b. Annual Access Conference
        i. This annual event has been held since 1993, and while it has a focus on information/library technology, it is one of the few conference events that features a hackfest, where teams gather around common problems, and develop solutions, including code.
4. International Event Bursaries
    a. Open Bioinformatics Foundation: OBF Fellowship

        i. Promotes open source bioinformatics software development and open science in the biological research community. The program is aimed at increasing diverse participation at events promoting open science practices such as resource development and dissemination in the bioinformatics and biological research community.These fellowships are available to support both in-person and remote (virtual) participation at events such as conferences, workshops, training courses, or collaborative development sprints. Events can include conferences; workshops, code fests, and hackathons. (Global)

Community Development

1. Open Source Program Offices (OSPOs)
    a. Alfred P. Sloan Foundation's [Call for Letters of Inquiry: Institutional Support for Open Source Software in Research]
        i. Grants of up to $750,000 over two years will be awarded to USA higher education institutions to launch university Open Source Program Offices (OSPOs)
2. RSE Events
    a. [SocRSE Events and Initiatives Grant]
        i. Financial support available for events and initiatives which support establishing a research environment that recognises the vital role of software in research. For example: increase software skills across everyone in research, to promote collaboration between researchers and software experts, and to support the creation of an academic career path for Research Software Engineers. Applications for support will follow a peer review process, this process happens monthly. (Global)
3. Promoting Community and Collaboration
    a. [Horizon Europe: Development of community-based approaches for ensuring and improving the quality of scientific software and code]
        i. Aims to promote quality of software and code across the different disciplines. Activities covered include those that foster alignment of existing initiatives by promoting coherence and developing community guidelines, and ensure integration of infrastructure, tools and services. (Global - at least one independent individual or legal entity established in a EU Member State)
4. RSE Community Groups
    a. In 2023 the US-RSE chapter received funding from the Alfred P. Sloan Foundation to "provide support in three important areas: the recruitment of dedicated staff, the development of activities and initiatives, and the enhancement of organizational tools and community well-being." Funding for the development of RSE Chapters internationally are growing, and highlight the importance of building community in the RS ecosystem.

Coordination

Access to RSEs

1. Short-Term (3-6 months)
    a. [New Zealand eScience Infrastructure Research Consultancy]
        i. Researchers from NeSI Collaborator institutions (University of Auckland, NIWA, University of Otago, and Manaaki Whenua - Landcare Research) and researchers with Merit projects can usually access Consultancy at no cost to themselves, based on their institution's or MBIE's investment into NeSI. Otherwise, please visit the Partners & Pricing page for costs related to our Consultancy service. The NeSI team are available to help with any stage of your research software development. We can get involved with designing and developing your software from scratch, or assist with improving software you have already written.
2. Medium-Term (6-12 months)
    a. Netherlands eScience Center [Call for Sustainable Software]
        i. The eScience Centre has a series of calls to provide applicants with access to in-kind RSE resources, with calls ranging from 3 months to 3 years. The example linked to above "supports communities of researchers who require their software to meet higher quality standards to ensure the continuity and advancement of their research in the longer term."
3. Long-Term (1-2 years)
    a. [Schmidt Futures Virtual Institute for Scientific Software (VISS)]
        i. VISS seeks to accelerate the pace of scientific discovery through the support and development of better quality, more sustainable scientific software. Starting with a network of four inaugural centers based at the University of Cambridge, Georgia Institute of Technology, the Johns Hopkins University, and the University of Washington, VISS will address the growing demand for high quality professional software engineers who can build dynamic, scalable, open software to facilitate accelerated scientific discovery across fields. The objective of these scientific software centers is to not only improve the quality of research and accelerate advancements, but to also support longer term platforms and systems that encourage best practice in open science. This will be achieved by providing scientific researchers with access to full-time professional engineers and state of the art technology and techniques such as high-end computing, massive databases, and machine learning.

HQP

- Institutional Teams
    a. [Schmidt Futures VISS Program] (see above - Access to RSEs, long-term)

b. CANARIE Local Research Software Development Teams
    i. This program "will support teams of three dedicated, full-time research software developers at each of the selected institutions. These teams will be available to support all researchers at the institution, regardless of discipline, and will provide guidance, training, expertise, and software development specific to advancing research projects." The program was considered a success, and led to the development of local RS teams that had institutional support once the CANARIE funding was expended.

- Regional Teams
  a. Alliance Support for Federation Staff
    i. The Alliance, as well as the predecessor programs from CFI that supported Compute Canada, provide support in various forms, including for staff at regional compute organizations. This support was most recently enhanced under the Alliance's MYFP 2023-25 mandate.
- Thematic Teams
  a. DFG/NFDI Support for Domain Consortia
    i. The NFDI Consortia represent 26 specific research domains, and while data-focused, include elements of RS and SW development as key aspects of the Consortia efforts. For example, the NFDI4Health Consortia has the following goal: "The NFDI is intended to systematically open up databases from science and research, secure them in the long term, make them accessible and network them (inter)nationally." Specific approaches include:
      1. Enable discoverability of and access to structured health data.
      2. Federal framework for data storage organizations received.
      3. Enable exchange and linking of personal data while maintaining data protection.
      4. Establish automated services (e.g., search, analysis tools).
      5. Establish and improve interoperability and reusability of data.
      6. Promote use case-oriented collaboration between research communities.
- RS Fellowships, Prizes and Training
  a. The eScience Fellowship Programme
    i. The program is aimed at members of the Dutch academic research community who are interested to act as ambassadors for the use of research software. Eligible candidates could be a researcher who focuses on the application of research software, or a developer who writes tools for researchers, or a research software engineer who supports the work of researchers with software, or an advocate for best practice in software use for their research domain, or an individual in a leadership role.
  b. Simons Foundation Scientific Software Research Faculty Award (SSRF)
    i. Funding to support new research professor positions in existing academic departments to be filled by scientific software-focused researchers. The

SSRF Award will support researchers who have a strong track record of leadership in scientific software development. The aim of this program is to stimulate the development and maintenance of core scientific software infrastructure in academic environments through creating a new, long-term, faculty-level career path.

   ii. 5 years of 50% salary support of the awardee's academic-year salary and fringe benefits, along with a yearly $50,000 research allowance for the awardee, as well as indirect costs for the host institution. The host university is expected to provide the other 50% of salary support for teaching work through existing department channels. This funding is potentially renewable.

c. Australian Research Data Commons (ARDC)'s Australian Museum Eureka Prize for Excellence in Research Software (and further details)
   i. This prize is awarded for the development, maintenance, or extension of software that has enabled significant new scientific research. The $10,000 Eureka Prize for Excellence in Research Software is open to both individuals and teams that include an Australian or Australia-based developer or maintainer. Entrants can either enter themselves or be nominated by others. (Australia)

d. European Master For High Performance Computing

EDI Focus

1. Access to Cyberinfrastructure
   a. National Science Foundation (NSF) Strengthening the Cyberinfrastructure Professionals Ecosystem:
      i. This program aims to democratize access to NSF's advanced cyberinfrastructure ecosystem; ensure fair and equitable access to resources, services, and expertise; by strengthening how Cyberinfrastructure Professionals function in this ecosystem.
2. Impacts
   a. The rOpenSci Champion Program pilot
      i. This call aims to identify, recognise, and reward passionate members of the R community, including people who belong to groups that are historically and systematically excluded. This is a 12-month program. (Global)

Appendix F: List of Abbreviations and Acronyms

Table 3. List of abbreviations and acronyms

Item	Spell-out	Item	Spell-out
LCDRI	Leadership Council for Digital Research Infrastructure	FRDR	Federated Research Data Repository
DM	data management	ARDC	Australian Research Data Commons
ARC	advanced research computing	BD2K	Big Data to Knowledge
ARS	advanced research software	NWO	Dutch Research Council
RS	research software	NeIC	Nordic e-Infrastructure Collaboration
RDM	research data management	UKRI	UK Research and Innovation
the RS Strategy	the Alliance RS Strategy for 2025-2030	NIH	National Institutes of Health
RSE	research software expert/engineer	CSSI	Cyberinfrastructure for Sustained Scientific Innovation
HQP	Highly qualified personnel	NLeSC	Netherlands eScience Center
RSSWG	RS Strategy Working Group	CZI	Chan Zuckerberg Initiative
ReSA	Research Software Alliance	DFG	German Research Foundation
EDIA	equity, diversity, inclusion, and accessibility	NSF	National Science Foundation
VREs	virtual research environments	IOF	Institutional Operating Funds
DOE	Department of Energy	CRSC	Canadian Research Software Workshop (later called Canadian

			Research Software Conference)
Federation	Alliance Federation (previously Compute Canada Federation)	TECC	Technical Experts of Compute Canada
HPCS	High Performance Computing Symposium	SocRSE	Society of Research Software Engineers
VISS	Virtual Institute for Scientific Software	LRSS	Local Research Software Support
CFI	Canada Foundation for Innovation	ACOA	Atlantic Canada Opportunities Agency
AIF	Atlantic Innovation Fund	OCE	Ontario Centres of Excellence
NSERC	Natural Sciences and Engineering Research Council	SSHRC	Social Sciences and Humanities Research Council
CIHR	Canadian Institutes of Health Research	OSPO	Open Source Program Office
SSI	Software Sustainability Institute	CANFAR	Canadian Advanced Network for Astronomical Research
CADC	Canadian Astronomy Data Centre	NRC	National Research Council
ONC	Ocean Networks Canada	EDI	Equity, Diversity, and Inclusion
IP	intellectual property	GPU	graphics processing unit
RSoF	Research Software Framework	CS	Current State
DS	Desired State	IT	information technology
FTE	full-time employee	CI/CD	Continuous Integration and Continuous Deployment

DRI	digital research infrastructure	HPC	High-performance computing
ML	Machine learning	AI	Artificial intelligence
RAC	Resource allocation competition	PaaS	Platform-as-a-Service
PI	Principal investigator	SOA	service oriented architecture
MCIN	McGill Centre for Integrative Neuroscience	HSS	Humanities & Social Sciences
IaaS	Infrastructure-as-a-Service	SaaS	Software-as-a-Service
sysadmin	system administrator	RAiD	Research activity identifier
RPaaS	Research Platform as a Service	CNN	convolutional neural network
MetRe	Metrics for Repositories	PID	Persistent identifier
RSECon	Research Software Engineering Conference	SSRF	Scientific Software Research Faculty Award
OEC	Open eScience Call	SS	Sustainable software
ASCR	Advanced Scientific Computing Research	ICGC	International Cancer Genome Consortium
IHMS	International Human Microbiome Standards	MME	Matchmaker exchange
IHEC	International Human Epigenome Consortium	HPOSS	High Priority Open-Source Science
R&D	research and development	NIWA	National Institute of Water and Atmospheric Research