

CARL and Portage Response to “Data Repository Selection: Criteria That Matter”

Ottawa, ON (January 25, 2021) – The Canadian Association of Research Libraries (CARL) and the Portage Network (Portage) would like to express concerns about the [Data Repository Selection: Criteria That Matter](#), recently proposed through the FAIR Sharing initiative. The proposed approach, developed primarily by a group of publishers, will have a negative effect on the development of Canada’s national infrastructure, on research equity, and on the rights of marginalized and under-served communities to fully participate in the scientific process.

Canada, like many other countries, is developing [national infrastructure and services](#) for research data management - including national, institutional and domain data repositories - that will support comprehensive research data management in this country. This is demonstrated, for instance, through the launch of [Dataverse Canada](#)¹, and the development of the [Federated Research Data Repository](#)².

We have several concerns about the data repository selection criteria:

National sovereignty and governance of Canadian research data

We believe it is in our national interest to take local responsibility for the stewardship of the valuable research data produced through Canadian research and to determine which criteria and practices are appropriate for repositories that collect Canadian-produced research data. A publisher-directed approach to data repository selection could result in a situation that is at odds with our national values and standards. We support the principles outlined in [UNESCO Recommendations on Open Science](#), that characterize open science as a global public good, stating “Open Science services should be viewed as essential research infrastructures, governed and owned by the community, and funded collectively by governments, funders and institutions reflecting the diverse interests and needs of the research community and society.”

¹ In partnership with the University of Toronto and Ontario Council of University Libraries Scholars Portal

² FRDR developed in partnership with the Compute Canada Federation, CARL, Portage, and several Canadian [partner organizations](#).

Publisher control of the repository market

The implementation of these “Data Selection Criteria” via a filtering tool for the selection of repositories would effectively provide their creators, mainly publishers, with a significant degree of control over the market for repositories, both by imposing conditions on existing repositories and controlling the entry of new repositories into the market. There is a significant risk that the implementation of the proposed criteria will become prescriptive and cede control over research infrastructure to a group of commercial, for-profit entities. Furthermore, the document does not address the governance of the criteria over time, leaving it unclear who will decide which repositories comply or how individual criteria will be determined.

Lack of consultation and transparency

While a stated aim of the proposed criteria is to support researchers, we note that researchers and research organizations appear to be absent from their development. As a matter of good practice, resources for a designated community should not be developed without input from that community. Researchers, research organizations, and the repository community should be consulted in the development of any recommended characteristics for repositories. Although comments were solicited in response to an earlier version of the document, there should be increased transparency about any feedback that has been received.

Diversity and equity

In [Canada's Science Vision](#), one of the expected outcomes is “Promoting equity and diversity in research”. The proposed selection criteria will limit repository options for researchers, undermining equity and diversity by concentrating data and decision-making in the hands of a small number of publishers and compliant repositories. By limiting researchers’ choices, the proposed criteria would have a negative effect on equity and diversity in research spaces. Researchers may be forced to deposit data in a particular publisher-endorsed repository as a prerequisite for publishing in a certain journal, rather than depositing their data in a more discipline-aligned, community-driven repository, or a national or institutional data repository.

Indigenous ownership of local knowledge

The principles of research data ownership, control, access, and possession (OCAP³), authored by the First Nations Information Governance Centre in Canada, underscore the importance of balancing global views on open science and equitable access with the rights and needs of Indigenous communities. The proposed criteria pose a risk to the rights of Indigenous Peoples with respect to ownership of their data. Contemporary research policies and best practices emphasize research partnerships that respect communities and promote their self-determination, sovereignty, autonomy, and rights of data ownership. [The First Nations Principles of OCAP[®]](#) and Article 31 of the [United Nations Declaration on the Rights of Indigenous Peoples](#) are expressions of these principles.

Conclusion

Canadian research data infrastructures are already engaged in several initiatives to achieve FAIR⁴, TRUSTworthy⁵, and CARE-compliant⁶ digital repositories. One such initiative is the [CoreTrustSeal](#) certification process for data repositories - a globally recognized, widely-adopted set of criteria, developed, endorsed, and managed by a broad range of RDM stakeholders. Looking forward, the data management community should aim toward the implementation of good practices, as determined by a range of stakeholders including domain communities, research funders and governments, Indigenous communities, repositories, and libraries. We believe that research data produced in Canada are a national public good, and that the range of repository choices available to Canadian researchers, and the associated features of those repositories, should be determined by the major RDM stakeholders in Canada.

ABOUT CARL: CARL provides leadership on behalf of Canada's 31 research libraries, enhancing their capacity to advance research and higher education. It promotes effective and sustainable knowledge creation, dissemination, and preservation, and advocates for public policy that enables broad access to scholarly information.

³ OCAP[®] - <https://fnigc.ca/ocap-training/>

⁴ <https://www.go-fair.org/fair-principles/>

⁵ <https://www.rd-alliance.org/trust-principles-trustworthy-data-repositories-%E2%80%93-update>

⁶ <https://www.gida-global.org/care>

ABOUT Portage: Portage was launched by CARL in 2015 to contribute to the shared stewardship of research data and to address specific gaps in national research data management (RDM) infrastructure in Canada. Initiated as a library-based network, Portage now collaborates with a broad range of stakeholders and partners locally, nationally, and internationally to build capacity and to develop services and infrastructure. Over 160 experts from 60 institutions across Canada actively participate in Portage committees and infrastructure projects, with coordination provided by a national secretariat.

Media contacts:

Susan Haigh
Executive Director, CARL
susan.haigh@carl-abrc.ca

Jeff Moon
Director, Portage Network
portagedir@carl-abrc.ca