NDRIO: Limit Corporate Influence, Maximize Public Involvement and Accountability

Catherine Stinson[1,2], LLana James[3], Mohamed Abdalla[4,5], Norma Möllers[6], Sharday Mosurinjohn[7], Susan Phillips[8,9], Amber L. Simpson[1,5,10]

[1]School of Computing, Queen's University, Kingston, ON
[2]Department of Philosophy, Queen's University, Kingston, ON
[3]Faculty of Medicine, University of Toronto, Toronto, ON
[4]Department of Computer Science, University of Toronto, Toronto, ON
[5]Vector Institute for Artificial Intelligence, Toronto, ON
[6]Department of Sociology, Queen's University, Kingston, ON
[7]School of Religion, Queen's University, Kingston, ON
[8]Department of Family Medicine, Faculty of Health Sciences, Queen's University, Kingston, ON
[9]Public Health Sciences, School of Medicine, Queen's University, Kingston, ON
[10]Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON

The Federal government's five-year investment of $375 million in the New Digital Research Infrastructure Organization (NDRIO), to advance and invest in national digital research infrastructure, is much welcomed and long awaited. NDRIO can ensure that Canadian researchers have the tools needed to engage in cutting edge research of service to the public. A public counterpoint to for profit data services is urgently needed given that major technology companies hold massive amounts of data about Canadian residents, have profits exceeding the GDP of most countries, avoid any significant contributions to our tax base, and sit poised to take over delivery of public services such as education, transportation and healthcare. However, NDRIO's call for white papers does not address this context. The call raises good questions, but the narrowness of the focus is concerning. Given that NDRIO is a national taxpayer-funded and not-for-profit organization, with a mandate to foster research that benefits society, the absence of questions framed in terms of the public interest is striking.

Two intertwined concerns are of particular urgency:

1. The danger of corporate influence on and control of research, and on decision-making about research infrastructure.
2. The absence of structures and processes to ensure full accountability to the public, that funds NDRIO and the majority of its member organizations, given the amount of power NDRIO will exercise with respect to data infrastructure, data management, re-use; and related innovation.

Section 1 of this white paper examines how NDRIO's questions for the research community prioritize individual/researcher interests and innovation, in alignment with corporate interests rather than public good.

Section 2 examines NDRIO's accountability to the public and poses questions for further consideration.

1. Limiting Corporate Influence

While starting a public consultation by calling for white papers is constructive and can be inclusive, the call focuses exclusively on technical issues that concern academics and data scientists. A truly public process should consult a wider spectrum of stakeholders, and invite comment not just on technical questions, but on privacy, fairness, and data justice. Public processes are inherently more likely to align with public interests, compared to a process directed at researchers and academics whose livelihoods sometimes depend on private funding, and thus have an embedded and inevitable conflict of interest. We note that NDRIO did not ask for white paper authors to submit statements of competing interests.

In a recent paper, Abdalla & Abdalla present worrying evidence about the prevalence of conflicts of interest among AI researchers, and the effects of Big Tech and industry funding on research. Across the four major computer science schools polled, more than half of AI faculty have funding from Big Tech. The paper argues that sponsorship is "actively distorting the academic landscape to suit its needs." Big Tech sponsorship includes funding major AI meetings (including meetings about AI Ethics); funding individual AI researchers; and major funding of AI institutes. NDRIO must limit conflict of interest in its operations and decision making. We make three recommendations in this vein:

1. Perceived, potential, current and previous conflicts of interests should be disclosed by the leaders, members, and users of public resources like NDRIO.
2. Private interests (be they represented by persons, partnerships, and/or organizations) should not have a seat at the table or have influence in decision-making or participate in decision making processes about publicly funded infrastructure.
3. In order to limit the capture of public resources by private corporations, access to data arising from non-commercial sources (like electronic medical records, the Ontario Health Data Platform, *etc.*) should not be made available to representatives of private/commercial enterprises under the same conditions of access as to independent researchers.

Disclosure of conflicts-of-interest is essential, because research done by private corporations is, by definition, motivated by profit. Corporate research has a long history of endangering the public. Some well-known examples include the lessons from corporate influence in the opioid epidemic, pharmaceutical industry selectively releasing only positive results of clinical trials, the tobacco industry suppressing research demonstrating links to lung cancer, oil companies sponsoring fringe climate change deniers, and, particularly pertinent to the DRI, Big Tech performing population-wide experiments on social media users for the purposes of manipulating human behaviour to increase profits. An especially egregious example of corporate control of which research questions get pursued happened recently with Google's firing of Timnit Gebru for pushing back against their attempt to censor a research paper produced by their in-house AI Ethics team that pointed out the harms (with respect to discrimination and environmental effects) of the large language models used in their products. Google's firing of Gebru illustrates the tension between corporate and public interests: as a private entity, Google has no obligation (nor mandate) to act in the public interest and publicize the environmental and social harms of their technologies. The lesson for NDRIO is that public infrastructure must be kept under public jurisdiction so that when public interests run up against profit motives, corporate misconduct is precluded, and the interests of the public are protected.

The public has little to gain from research performed by private corporations, when the IP of research using non-commercial data (e.g., health data) does not belong to the general public. Private corporations often leverage their research and IP to obtain "excessive monopoly power" in their provision of commercial services in healthcare, transportation, and other formerly public services, resulting in poorer services for the public. An example is the quashed plan of Sidewalk Labs to build civic infrastructure and levy taxes in Toronto's Waterfront, while conducting an unprecedented experiment in ubiquitous surveillance. This plan demonstrated the unbridled corporate aspirations of capturing Canadian public resources, as well as the size of the public reaction needed to curtail them. What public reassurance is there that NDRIO will follow the public mandate of adopting human-centric AI grounded in human rights, rather than conforming to industry standards? Accountability measures need to be put in place and communicated publicly.

Here, NDRIO can improve upon past federal investments in, for example, Toronto's Vector Institute, where no such accountability measures seem to be in place. Research talks by Vector Institute members are not open to the public, nor even to their network of Affiliate Members. As Abdalla & Abdalla point out, Vector Institute

faculty members include(d) "a Vice President of Google, and the heads of various companies' AI branches such as Uber, and Nvidia," and a third of their funding comes from industry. Although a few members are prominent AI Fairness researchers, their work centres only on defining mathematical formalizations of fairness. While originally intended as a tool for fighting discrimination, this kind of fairness research has been co-opted by Big Tech as a tool that lends the appearance of ethical self-regulation while not making a meaningful difference, a practice known as "ethics washing". If NDRIO is to serve the public interest, then it must ensure public accountability with respect to human-centric AI grounded in human rights, fairness, equity, transparency and beneficence.

2. For the Public Good

The opportunity for, and urgency to establish public accountability mechanisms is especially acute in the case of the DRI. This is not only because the DRI is intended to benefit Canadian residents, but also because it *contains* personal data about Canadians - including highly sensitive data such as health information. Communities who are meant to benefit from public goods should have a voice in how research is to be conducted. If no such mechanisms for community participation are implemented, NDRIO will position researchers in a fundamentally exploitative relationship with the public, a relationship in which the population serves as mere data fodder for research without any power to hold scientists accountable.

A wealth of empirical research has demonstrated that data science and AI inherit and exacerbate the problems of underlying skewed (e.g. racist) data sets. Given this evidence of discriminatory effects of AI and contemporary data science across a range of public institutions -- from healthcare to education to the criminal justice system - it is especially important that NDRIO develops clear and convincing harm-prevention measures. Canada also has a special responsibility to ensure that research questions, data collection, analysis and interpretation do not add to historic harm to Indigenous populations or individuals. Canada has a special responsibility to ensure that colonial and racist data collection about marginalized communities does not continue to harm these communities. Thus, NDRIO needs to not only demonstrate how the DRI is going to benefit the public, but develop harm-prevention mechanisms, which may, for example, include regular audits that control for harm done to communities which have been historically marginalized.

DRI's stakeholders need to include individual members of the public and academics with no competing interests rather than corporations. To have credibility NDRIO should develop clear accountability mechanisms that ensure the involvement of Canadian residents and communities in the DRI's governance and in all stages of research and procurement decisions. We suggest that NDRIO consult publicly to develop answers about accountability in the following three areas:

1. Research for the public good (who benefits from the DRI?):
   ○ How will NDRIO ensure that the communities who are intended to benefit from related research will shape the direction and conduct of the research?
   ○ What procedures will ensure that those most affected by a study are engaged and heard?
   ○ How will NDRIO ensure that corporations do no access, misuse or re-identify sensitive data (e.g. health data)?

2. Managing and protecting individuals' data:
   ○ What are the procedures to ensure that the context in which data were collected is recorded and provided with the data?
   ○ What procedures will ensure scrutiny to identify biased or skewed information in the data itself?

- How will NDRIO ensure informed consent for every new project or for shifts in scope of a previous project by the people who donated their data? Is informed consent possible as technology becomes increasingly complex?
- How can NDRIO work towards transparency of data so that communities can see what data are collected about them?

3. <u>Harm-prevention mechanisms:</u>
    - How will NDRIO demonstrate accountability for possible harms arising from data use?
    - What restitution and sanction processes will be in place to mediate conflicts in case harm is done such as if research or technology is found to be discriminatory?
    - What are the procedures to make sure that data and analysis can be corrected or deleted when bias is identified?