

Developing Big Data Analysis for Public Benefit

Lynn Eakin, Ontario Nonprofit Network

Big data is a term for “datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze.” [1] The world now has the capacity to use big data for a range of applications.

In the nonprofit sector, as well as within government, we are in the early stages of exploring the uses of big data for our work in communities. Both government and the nonprofit sector have data – lots of data – but we are just at the start of being able to harness this data and use it to improve our work. This is why the open data movement is so important.

How can we use government and sector data, along with other data, to do our work better? How can big data assist us to tackle the challenges we face – poverty, growing inequality, climate change, physical and mental health, unemployment and underemployment, homelessness, and food security?

In businesses, big data has been used with success in commerce, for example: in helping to identify and match products with customers by aggregating information on consumers to better understand their purchasing patterns, lifestyles, etc. Big data is also used in other areas such as in the investment field and to curate music and films.

Big data has the potential to inform the nonprofit and government sectors. It is already being used in the health care system to inform best practices, funding decisions and treatment availability. However, the community-based nonprofit sector and much of government is lagging in its ability to open its data and to analyze that data for public benefit.

We recognize that the nonprofit sector, government and private sector need to partner together to develop our collective capacity for big data use and analysis.

We need to identify what data is useful for our work and explore together the possibilities of big data to inform, streamline and potentially transform how we do our work together.

While everyone agrees that ensuring the well being of our citizens and building strong resilient communities is the shared goal, the use of big data is not bias free. It is not merely a question of the aggregation and open publication of data.

The initial collection and analysis of this data requires decisions based on the curated knowledge of experts, community values and priorities. These factors are critical as big data is only one piece of the puzzle. For discussion of the challenges of using big data to address social problems the big data for good discussion blog is a good place to start. [2]

Decisions regarding the use of data sets, the emphasis and value placed on different criteria and the accuracy of the findings cannot simply be left to proprietary interests with private financial not public benefit objectives.

It is essential, therefore, that public sector big data (sector and government data) not only be published in the public domain, but that the data analysis and methodology must be open and include key stakeholders in its development.

"Big data has the potential to inform the nonprofit and government sectors. However, the community-based nonprofit sector and much of government is lagging in its ability to open its data and to analyze that data for public benefit"

Decisions regarding the use of data sets, the emphasis and value placed on different criteria and the accuracy of the findings cannot simply be left to proprietary interests with private financial not public benefit objectives.

It is essential, therefore, that public sector big data (sector and government data) not only be published in the public domain, but that the data analysis and methodology must be open and include key stakeholders in its development.

The following principles can guide big data development for the public domain:

- The default position on data in the nonprofit and government sectors should be open. Work needs to be undertaken to develop standards to make this data readable and accessible for use and to assist nonprofit and government data holders open up their data.
- Big data use and analysis for public good purposes needs to be publicly owned and open. A nonprofit consortium of government, nonprofit sector, funders, academics/universities, private sector and data managers should develop big data mechanisms for the public domain.
- The intellectual property generated from big data in the public domain must be held for the public good. Intellectual property generated in the development of big data for public purposes and public benefit cannot be held in proprietary ownership.
- Social innovation is complex and big data will be only one of many important contributors to finding new ways forward. Values, priorities, and local and individual circumstances will continue to play a significant role.

Notes:

[1] <http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation>

[2] <http://www.odbms.org/blog/2012/06/big-data-for-good/>

Data matters. It helps us figure out how communities and people are doing. But there is a shortage of high-quality, relevant, and current data on the nonprofit labour force and economy in Ontario and Canada.

The shortage of sector data makes it difficult to propose sector-wide solutions and to describe the sector's economic and societal impacts.

ONN is concerned about three issues around data:

Shortage of data. Governments, academics and researchers gather a lot of detailed information on labour markets and the economic impact of many sectors of Canada's economy. But not on the nonprofit sector. The two main ways data is classified in this country is by industry and by job type. But those classifications don't make clear the sector – so we don't know if someone works in the nonprofit, for-profit or public sector. For example, computer engineers could work in any of the three sectors; and organizations falling under the industrial code of 'retail trade' 'professional services', 'food services' and more could be operating as a for-profit or a nonprofit.

Out-of-date nonprofit data: Out-of-date nonprofit data: Data that is specific to the nonprofit labour force consists of national surveys that are out of date. While Statistics Canada released new composite data on the nonprofit sector's economic contribution in March 2019 after years of advocacy on the part of our sector, there is no commitment for ongoing data collection or a dedicated sector survey. And smaller-scale surveys that are not necessarily representative of the sector as a whole.

Lack of data sharing. Nonprofits and governments collect a great deal of operational, or administrative, data on the people they serve. But they don't share this data. This is a missed opportunity. We think there could be great public benefit if nonprofits and governments figure out a confidential, responsible, standardized, way to share that data. Sharing could promote new opportunities for measuring program outcomes, providing more collaborative care, and supporting evidence-based policy.

Read more: theonnc.ca/our-work/our-people/data-strategy