

Data collected by governments can be useful to researchers, but only when accessed carefully

1 August 2019, by Jack Teng



Credit: CC0 Public Domain

Data are generated every time we make a purchase or receive other services such as health care. This has always been the case, but over the past 20 years, data collection has become increasingly automated, with data collected and stored in digital (rather than paper) formats.

Data collected during the provision of government services are known as administrative data (for example, applying for a driver's license or updating health card information). These data are particularly powerful for researchers because they contain a wide range of information about whole populations. These [large data sets](#) can then be studied to derive important findings that are [clinically or socially meaningful](#).

Digital storage increases the potential for the use of data, and increased computing power makes it easier to study the data and find actionable

insights. Personal data are being used to make decisions in the public and private sectors. Public concerns over how data are used, stored and who has access to them is forcing [government agencies](#) to take a new look at the data they collect and what they do with it.

Our recent research, published in the *International Journal of Population Data Science*, [reports the recommendations made by 24 members of the public during a public deliberation](#) —a kind of public engagement event. This event was held in Vancouver, B.C., and was hosted by [Population Data BC](#), an organization at the University of British Columbia that facilitates data sharing for research and provides education and training on data use.

Digital information production

There are both potential risks and benefits to the digital society in which we currently live. The risks include [possibilities for surveillance, loss of privacy, discrimination and loss of reputation and autonomy](#).

On the other hand, benefits include the use of data for research, leading to new insights about health, diseases and effective services [to promote and support the flourishing of individuals and communities](#). How data are used may also have important ramifications for particular groups of people or communities, as well as individuals.

For example, an analysis of data might inform decisions about [health-care](#) delivery; for example, a decision that group X is not likely to benefit from a particular service and therefore should not be eligible for that service. In fact, when public impacts such as these have not been considered or when the public has not been consulted about them, [data sharing projects may fail](#).

There are strict government rules surrounding how [administrative data can be shared with researchers](#). These rules protect data, sometimes at the expense of research, [when they make data either slow or impossible to acquire](#).

Managing access to data

Most forms of data are not held together in one huge database in one location, nor are they controlled by one organization. Instead, each agency—for example, a health ministry—or a department within a large agency, manages its data.

In most cases, there is a designated data steward or privacy officer who approves or denies data access requests. Each individual may have different interpretations of privacy regulations or they may disagree on what to share with the researchers; this can result in delays for the researchers to receive the data and start their research.

Privacy legislation and other rules help [set some boundaries for data use](#), but the data landscape is changing rapidly as academic and government researchers require for greater access. The big questions then are: What role do we want big data to have in our society? How can the risks and benefits of data use be balanced?

There is no obvious answer to these questions as they involve consideration of conflicting values such as differences between the importance of maintaining privacy and the importance of conducting research for the public benefit. The best balance may be assessed differently by different groups and individuals. As such, it is important to involve the public in establishing norms and guidance for [data use](#).

Public input on data management

Public deliberations are distinguished from surveys, focus groups or other types of engagement by their focus on producing [collective and civic-minded responses to policy issues](#). This requires time for the participants to meet and engage with each other, receive and understand technical information

and understand the diversity of social perspectives with regard to their opinions about data sharing.

In our study, participants received an information booklet and then met for four days over two non-consecutive weekends. During this time, they heard presentations from five experts on various topics related to data sharing such as privacy legislation, patient concerns and researcher data needs. Participants then deliberated on the issues. Participants were selected to reflect B.C. demographics in terms of factors such as age, ethnicity, income and geographic location, as much as possible.

Participants developed and voted on [19 recommendations for consideration by policy makers](#) to improve B.C.'s existing data sharing system. These recommendations started by acknowledging clear support for research uses of administrative data and an equal expectation that data would be kept safe and secure both in general and during the research process.

Participants also wanted the data sharing system to be efficient. Several of their recommendations suggested possible improvements in system efficiencies, including standardizing the policies and procedures that data stewards follow when assessing data sharing requests. They suggested different ways of achieving this, such as developing training and certification programs for data stewards.

Participants made it clear that research is important, but that does not mean that all research requests should be automatically granted. Researchers were seen to have responsibilities as well, including transparency about how the data are collected and used, especially with the communities and vulnerable populations from which the data derive.

Clear support for administrative data for research in this deliberation is consistent with [other similar public deliberations around the world](#). Participants' recommendations provide helpful insight into what is important to them when it comes to sharing data—this type of information is critical to help design data sharing systems that are acceptable and

worthy of public trust.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

APA citation: Data collected by governments can be useful to researchers, but only when accessed carefully (2019, August 1) retrieved 21 October 2019 from <https://techxplore.com/news/2019-08-accessed-carefully.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.