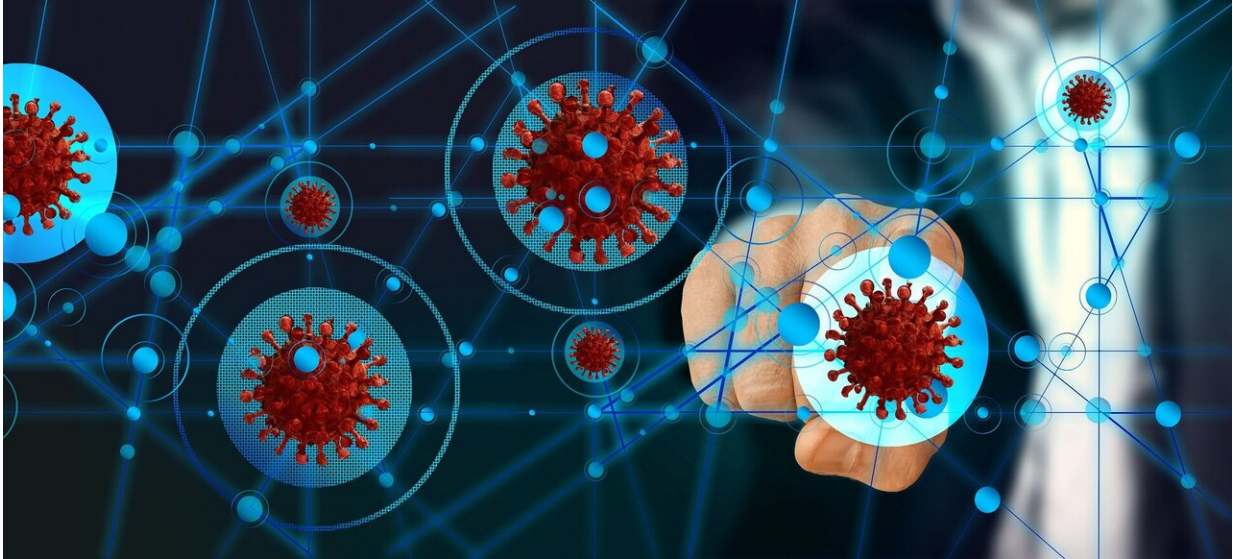# Video: Who controls pandemic data?

August 3 2020, by Julia Lane

Credit: Pixabay/CC0 Public Domain

Editor's note: When the Trump administration ordered hospitals to report COVID-19 data to the Department of Health and Human Services rather than the Centers for Disease Control and Prevention as they had been doing, it provoked worries and criticism from public health experts. The White House said that the HHS system will provide more accurate data faster, but the switch did raise concerns that political considerations would influence what data is reported. Professor of public policy Julia Lane, who recently published the book "Democratizing Our Data: A Manifesto," explains why public data is vital to public health and democracy in general.

What was the main concern over the data?

The whole point of having a career civil service running public data systems is that, because they can't be fired, they have the integrity to produce the statistics the best way possible. And that's what makes the [federal government](#) and state and [local governments](#) such high-quality data engines.

Now, the concern that came up is the appearance of political interference. Who knows what actually happened. But the point is, if there is political pressure on the measurement, then that can substantially affect the aggregates. The language that has come out of the administration has not helped the cause of the career civil servants appropriately.

Why is it important to have accurate and transparent public data?

When you're making decisions that are important for all the citizens of the country, or the population of the country as a whole, then you need good data to be able to allocate those resources. Now, if those data are biased in some way, people are not going to get counted. And if they're not counted, they're not going to get resources.

People matter. A democracy is a government of the people, by the people and for the people. If you don't know who the people are, you don't have a democratic system of government. And if you don't have high quality data, you can make lots of mistakes. For example, we didn't have high quality data on the [opioid crisis](#). And so it kind of surprised everyone how bad it was because we had no way of measuring it.

What happens when government data is influenced by politics?

In the United States, I don't think that has been a major issue, although I'll give you one example in which government data has been influenced by politics. But certainly in dictatorships, government data is influenced by politics because if you control the message of the data, you control an awful lot of messaging that's going on in the country. Anyone who's worked for the World Bank or in totalitarian countries will be able to tell you that [government](#) data is the first thing that goes.

Now, I'm going to give you an example from the United States, and this is quite well documented. So the U.S. Census Bureau in 1940 was asked to provide tabular information on the location of Japanese Americans. [That's the information that was used](#) to round people up and put them in internment camps—Japanese Americans in internment camps.

People are relying on nongovernmental sources, such as like Johns Hopkins University or the media, for data on the spread of the virus. What are some potential problems with data from private institutions?

The challenge is if you don't have a trusted source and what you're seeing happening here is people are going to multiple other sources. So they're going to [Johns Hopkins](#), [Worldometer](#) or [1Point3Acres](#) – people are getting their data from lots of different sources.

I don't want to cast aspersions on any of those datasets, but how does the data that they put out compare with some measure of ground truth? How does the data collection persist over time? How do we standardize measures across countries? With private institutions, maybe people are trying to sell you things. Maybe there's marketing involved or there's a profit motive.

How do we improve our public data systems?

What I talk about in the book, which is called "Democratizing Our Data: A Manifesto," is reducing the monopoly power in the federal system. If you have a monopoly power, you've got a single point of failure, and that makes you vulnerable to these political pressures that we're seeing.

So what I talk about is a networked system that pushes the development of measures and indicators down to the states and local areas—the regions which are closer to the data and have a better sense of the way in which the data are generated. But combine that with the federal system so that you get consistency, that quality focus that I was just talking about.

The current system clearly isn't working. When I wrote the book, I did not expect the coronavirus pandemic to highlight all of the fragilities in our data collection system. I talk much more about GDP and unemployment. But all of the fragilities of our current system are being exposed with the COVID-19 pandemic.

This article is republished from The Conversation under a Creative Commons license. Read the original article.

Provided by The Conversation