



GarryKillian/Shutterstock.com

Six ways (and counting) that big data systems are harming society

December 7, 2017 10.32am GMT

There is growing consensus that with big data comes great opportunity, but also great risk.

But these risks are not getting enough political and public attention. One way to better appreciate the risks that come with our big data future is to consider how people are already being negatively affected by uses of it. At Cardiff University's **Data Justice Lab**, we decided to record the harms that big data uses have already caused, pulling together concrete examples of harm that have been referenced in previous work so that we might gain a better big picture appreciation of where we are heading.

We did so in the hope that such a record will generate more debate and intervention from the public into the kind of big data society, and future we want. The following examples are a condensed version of our recently published **Data Harm Record**, a running record, to be updated as we learn about more cases.

Author



Joanna Redden

Lecturer in Critical Data Studies, Co-Director Data Justice Lab, Cardiff University

1. Targeting based on vulnerability

With big data comes new ways to socially sort with increasing precision. By combining multiple forms of data sets, a lot can be learned. This has been called “algorithmic profiling” and raises concerns about how little people know about how their data is collected as they search, communicate, buy, visit sites, travel, and so on.

Much of this sorting goes under the radar, although the practices of data brokers have been getting attention. In her testimony to the US Congress, World Privacy Forum’s Pam Dixon reported finding data brokers selling lists of rape victims, addresses of domestic violence shelters, sufferers of genetic diseases, sufferers of addiction and more.

2. Misuse of personal information

Concerns have been raised about how credit card companies are using personal details like where someone shops or whether or not they have paid for marriage counselling to set rates and limits. One study details the case of a man who found his credit rating reduced because American Express determined that others who shopped where he shopped had a poor repayment history.

This event, in 2008, was an early big data example of “creditworthiness by association” and is linked to ongoing practices of determining value or trustworthiness by drawing on big data to make predictions about people.

3. Discrimination

As corporations, government bodies and others make use of big data, it is key to know that discrimination can and is happening – both unintentionally and intentionally. This can happen as algorithmically driven systems offer, deny or mediate access to services or opportunities to people differently.

Some are raising concerns about how new uses of big data may negatively influence people’s abilities get housing or insurance – or to access education or get a job. A 2017 investigation by ProPublica and Consumer Reports showed that minority neighbourhoods pay more for car insurance than white neighbourhoods with the same risk levels. ProPublica also shows how new prediction tools used in courtrooms for sentencing and bonds “are biased against blacks”. Others raise concerns about how big data processes make it easier to target particular groups and discriminate against them.

And there are numerous reports of facial recognition systems that have problems identifying people who are not white. As argued here, this issue becomes increasingly important as facial recognition tools are adopted by government agencies, police and security systems.





Facial recognition. Zapp2Photo/Shutterstock.com

This kind of discrimination is not limited to skin colour. One study of Google ads found that men and women are being shown different job adverts, with men receiving ads for higher paying jobs more often. And data scientist Cathy O’Neil has raised concerns about how the personality tests and automated systems used by companies to sort through job applications may be using health information to disqualify certain applicants based on their history.

There are also concerns that the use of crime prediction software can lead to the **over-monitoring of poor communities**, as O’Neil also found. The inclusion of nuisance crimes such as vagrancy in crime prediction models distorts the analysis and “creates a pernicious feedback loop” by drawing more police into the areas where there is likely to be vagrancy. This leads to more punishment and recorded crimes in these areas.

4. Data breaches

There are numerous examples of data breaches in recent years. These can lead to identity theft, blackmail, reputation damage and distress. They can also create a lot of anxiety about future effects. One study discusses these issues and points to several examples:

- The Office of Policy Management breach in Washington in 2015 leaked people’s fingerprints, background check information, and analysis of security risks.
- In 2015 Ashley Madison, a commercial website billed as enabling extramarital affairs, was breached and more than 25 gigabytes of company data including user details were leaked.
- The 2013 Target breach in the US resulted in leaked credit card information, bank account numbers and other financial data.

5. Political manipulation and social harm

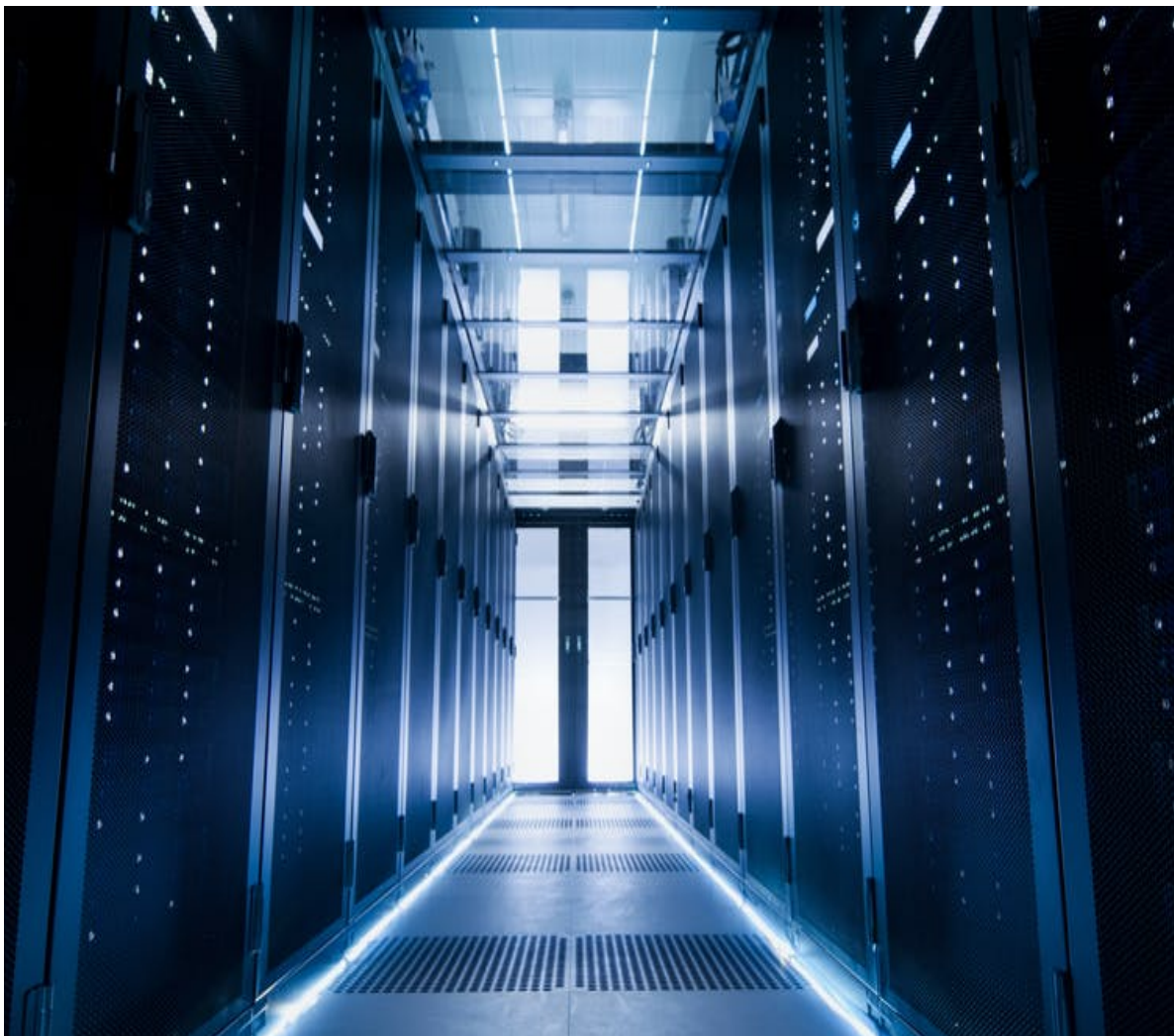
Fake news, bots and filter bubbles have been in the news a lot lately. They can lead to social and political harm as the information that informs citizens is manipulated, potentially leading to misinformation and undermining democratic and political processes as well as social well-being.

One recent study by researchers at the Oxford Internet Institute details the diverse ways that people are trying to use social media to manipulate public opinion across nine countries.

6. Data and system errors

Big data blacklisting and watch-lists in the US have wrongfully identified individuals. It has been found that being wrongfully identified in this case can negatively affect employment, ability to travel – and in some cases lead to wrongful detention and deportation.


In Australia, for example, there have been investigations into the government's automated debt recovery system after numerous complaints of errors and unfair targeting of vulnerable people. And American academic Virginia Eubanks has detailed the system failures that devastated the lives of many in Indiana, Florida and Texas at great cost to taxpayers. The automated system errors led to people losing access to their Medicaid, food stamps and benefits.



Data stored in centres such as this isn't necessarily safe. Gorodenkoff/Shutterstock.com

We need to learn from these harms. There are a range of individuals and groups developing ideas about how data harms can be prevented. Researchers, civil society organisations, government bodies and activists have all, in different ways, identified the need for greater transparency, accountability, systems of oversight and due process, and the means for citizens to interrogate and intervene in the big data processes that affect them.

What is needed is the public pressure and the political will and effort to ensure this happens.

 [Discrimination](#) [Big data](#) [Data](#) [Gender discrimination](#) [Racial discrimination](#) [Class divides](#) [Misinformation](#)
[Fake news](#) [Interdisciplinarity](#) [Data breaches](#)