The Harm That Data Do

Paying attention to how algorithmic systems impact marginalized people worldwide is key to a just and equitable future

By Joanna Redden | Scientific American November 2018 Issue

In Australia, they call it "robo debt": an automated debt-recovery system that is generating fear, anxiety, anger and shame among those who rely on, or who have relied on, social support. In 2016 the country's Department of Human Services introduced a new way of calculating the annual earnings of welfare beneficiaries and began dispatching automated debt-collection letters to those identified as having been overpaid. The new accounting method meant that fortnightly income could be averaged to estimate the income for an entire year—a problem for those with contract, part-time or precarious work. Reports indicate that the system went from sending out 20,000 debt collection notices a year to sending up to that many every week.

Previously, when the system identified someone who may have been overpaid benefits, a human was tasked with investigating the case. Under the automated system, however, this step was removed; instead it became the responsibility of the recipients to prove that they had not. That meant finding out why they were targeted—often requiring hours on the phone—and digging up copies of pay slips from as far back as seven years. To make matters worse, many of the debt notices were sent to people already living in precarious situations. Those targeted felt powerless because they had little time or resources to challenge the system. Newspapers reported at least one suicide. A social service organization would eventually report that a quarter of the debt notices it investigated were wrong, and an Australian senate inquiry concluded that "a fundamental lack of procedural fairness" ran through the entire process.

We have entered an "age of datafication" as businesses and governments around the world access new kinds of information, link up their data sets, and make greater use of algorithms and artificial intelligence to gain unprecedented insights and make faster and purportedly more efficient decisions. We do not yet know all the implications. The staggering amount of information available about each of us, combined with new computing power, does, however, mean that we become infinitely knowable—while having little ability to interrogate and challenge how our data are being used.

At the Data Justice Lab at Cardiff University in Wales, we maintain a Data Harm Record, a running log of problems with automated and algorithmic systems being reported from across the globe. We analyze this record to understand the diverse ways in which such systems are going wrong, how citizen's groups are dealing with the emerging problems, and how agencies and legal systems are responding to their challenges. Our studies, we hope, will result in a deeper understanding of how democratic institutions may need to evolve to better protect people—in particular, the marginalized—in the age of big data.

DEEPENING INEQUALITY

The robo-debt scandal is one of many that demonstrate the power imbalance incorporated into many emerging data systems. To understand what happened, we need answers to such

questions as why a system with such high error rates was introduced without adequate due process protections for citizens, why robust impact assessments were not done before it was rolled out, why the needs of those affected were not fully considered in designing the online portal or helpline, and why it was deemed permissible to remove human oversight. The problems with it, and with many other data-driven systems, stem in significant part from underlying social and political contexts: in particular, long-standing binaries of "deserving" and "undeserving" citizens that influence how they are valued and treated.

Some amount of error is inevitable in automated systems, as mathematician Simon Williams of Flinders University in Australia pointed out with regard to the robo-debt case: there always will be false positives and false negatives. Even so, seemingly random mistakes sometimes turn out to be discriminatory in nature. For example, face-recognition technologies routinely fail to identify nonwhite faces—which is a problem when that influences your ability to travel or to access government services. Joy Buolamwini, founder of the Algorithmic Justice League at the M.I.T. Media Lab, argues that this happens, in part, because the machine-learning algorithms are trained on data sets containing mainly white faces. Employees at the high-tech firms that designed these systems are, for the most part, white—an imbalance that can limit the ability to spot and address bias.

Likewise, an investigation by the research group ProPublica discovered that algorithms predicting the likelihood that someone charged with a crime would reoffend were twice as likely to falsely rank black defendants as high risk than white defendants. Similar scoring systems are being used across the U.S. and can influence sentencing, bonds and opportunities to access rehabilitation instead of jail. Because the models are proprietary, it is difficult to know why this happens, but it seems to be connected to weights the algorithms assign to factors such as employment, poverty and family history. Data drawn from a world that is unequal will reflect that inequality and all too often end up reinforcing it.

Disturbingly, researchers find that those at the top—the designers and the administrators—routinely fail to appreciate the limitations of the systems they are introducing. So, for example, the underlying data sets might contain errors, or they could have been compiled from other data sets that are not particularly compatible. Often, too, the implementers are unaware of bureaucratic or infrastructural complexities that can cause problems on the ground. They routinely fail to assess the impact of the new systems on the marginalized or to consult with those who do have the necessary experience and knowledge. When algorithms replace human discretion, they eliminate corrective feedback from those affected, thereby compounding the problem.

At other times, harm results from the way big data are used. Our data "exhaust"—meaning, the data we emit as we communicate online, travel and make transactions—can be combined with other data sets to construct intimate profiles about us and to sort and target us. People can be identified by religion, sexual preferences, illnesses, financial vulnerability, and much more. For example, World Privacy Forum's Pam Dixon found data brokers (the companies that aggregate and sell consumer data) offering a range of problematic lists, such as of individuals suffering from addictive behavior or dementia and rape victims, among others. Researchers studying the financial crash of 2008 found that banks had combined offline and online data to categorize and influence customers. In 2012 the U.S. Department of Justice reached a $175-million settlement with Wells Fargo over allegations that it had systematically pushed African-American and Hispanic borrowers into more costly loans.

Overall, the kinds of damage that data systems can cause are incredibly diverse. These may include loss of privacy from data breaches, physical injury as workplace surveillance compels people to do more than they can, increased insurance and interest rates, and loss of access to essentials such as food, home care and health care. In unequal societies, they serve to further embed social and historical discrimination.

DISSENT AS NECESSITY

What happens when people try to challenge data harms? To date, we have investigated cases involving governmental use of new data systems in Australia, Canada, the Netherlands, New Zealand, the U.K. and the U.S. Even in these democratic societies, relying on legal systems alone is usually not enough. Citizens are trying to obtain a measure of redress by combining their time and other resources into a collective and multipronged effort that includes all the pillars of democracy.

In the robo-debt case, those affected created a Not My Debt campaign for publishing their stories anonymously, getting help and sharing resources. According to Victoria Legal Aid senior executive Dan Nicholson, the organization has yet to initiate a federal court challenge, in part because people are reluctant to go public after the Department of Human Services released the private details of one critic to the press. One of his biggest concerns is how the government shifted responsibility to individual citizens for proving that no debt is owed, despite its vastly superior ability to compile evidence. The department says it has made changes to the system in response to early critiques, but experts say these are not enough. In its last budget the Australian government announced plans to expand the program.

In the Netherlands, individuals and organizations came together to launch a district court challenge against the government over Systeem Risico Indicatie (SyRI), which links citizen data to predict who is likely to commit fraud. The litigants argue that the system violates citizens' rights by treating everyone as guilty until proven innocent. The outcome of this court case will likely inspire citizens in other democracies seeking to protect their rights and to expand the definitions of harm. In the U.K., groups such as defenddigitalme are raising concerns about the psychological and social impact of Web-monitoring software in schools and the ways it can damage students who are wrongly labeled, for instance, as being suicidal or as gang members. In New Zealand, nongovernmental organizations (NGOs) successfully blocked an attempt by the Ministry of Social Development to require all providers of social services to provide data about their clients to receive government funding. The NGOs argued that the requirement could prompt members of already marginalized groups, such as refugees or victims of domestic violence, to avoid help for fear of being identified.

In Little Rock, Ark., an algorithm introduced by the state's Department of Human Services was blamed for unjustly cutting the home care hours of people with severe disabilities. Earlier, home care nurses determined home care hours. Now they helped people fill out a questionnaire and entered the data into a computer system—but it was the algorithm that decided. Government representatives argue that the automated system ensures that assignments of home care hours are fair and objective. Some individuals strongly disagreed, and with the help of Legal Aid of Arkansas, seven of them took the department to court. Six had seen their weekly home care hours cut by more than 30 percent. Court documents make grim reading, as each plaintiff recounts the impact of the cuts on their lives and health.

Examining information about the algorithm extracted via a court order, Legal Aid of Arkansas lawyer Kevin De Liban found numerous problems with it and how it was implemented. In May a judge ordered the Department of Human Services to stop using it, but the agency refused—whereupon the judge found the department in contempt. As the legal battle continues, the quality of life of thousands hangs in the balance.

These cases speak to the importance of collective mobilization in protecting people from injustices committed via data systems. It is difficult for individuals, in particular if they belong to marginalized groups, to interrogate the systems alone or to seek redress when they are harmed. Apart from instigating collective challenges, broader public discussion is needed about the transparency, accountability and oversight of data systems required for protecting citizens' rights. Further, how should information about these new systems be communicated so that we all understand? What are governments' obligations to ensure data literacy? And are there no-go areas? Surely maps of where and how governments are introducing data systems and sharing people's data should, as a first step, be provided as a matter of democratic accountability.

Just as important is ensuring that citizens can meaningfully challenge the systems that affect them. Given that datafied systems will always be error-prone, human feedback becomes essential. Critiques should be welcomed rather than fended off. A fundamental rethink of governance is in order—in particular, on the question of how dissent and collaboration might be better fostered by public bodies and authorities in societies permeated by data.