collective wellbeing Carnegie UK

Data Justice Lab

Western FIMS

Automating Public Services:

Learning from Cancelled Systems

Joanna Redden, Jessica Brand, Ina Sander and Harry Warne

Data Justice Lab

Acknowlegements

Thank you to the experts who generously gave their time by speaking with us, sharing information online and/or providing feedback at various stages of this project.

Neil Ballantyne, Open Polytechnic of New Zealand
Andrew Chen, University of Auckland
Lina Dencik, Data Justice Lab
Jamie Garcia, Stop LAPD Spying Coalition
Dominik Gerstner, Max-Planck Institute for the Study of Crime, Security and Law (MPI)
Merel Hendrickx, Public Interest Litigation Project (PILP-NJCM)
Arne Hintz, Data Justice Lab
Ronald Huissen: Secretary for Platform for Civil Rights
Fieke Jansen, Data Justice Lab
Thomas Berlin Hovmand, Director of education and culture, Gladsaxe municipality
Nicolas Kayser-Bril, AlgorithmWatch
Emily Keddell, University of Otago
Hamid Khan, Stop LAPD Spying Coalition
Cynthia Khoo, University of Toronto
Susanne Krasmann, Hamburg University
Susan Lindsay, Law Commission of Ontario
Jennifer Lord, PitMcGhee Palmers and Rivers
Camelia Naguib, Assistant Inspector General, LAPD Office of the Inspector General
Dan Nicholson, Victoria Legal Aid
Jędrzej Niklas, Data Justice Lab
Rashida Richardson, Northeastern University
Kate Robertson, CitizenLab
John Scott, QC
Luke Shaefer, University of Michigan
Rachel Siewert Senator, Australian Greens
Mackenzie Smith, Radio New Zealand
Yolanda Song, University of Toronto
Emiliano Treré, Data Justice Lab
Aneurin Thomas, Law Commission of Ontario
Joel Townsend, Victoria Legal Aid
Laura Tribe, OpenMedia Canada
William Webster, Centre for Research into Information, Surveillance & Privacy
Tijmen Wisman, University of Amsterdam
Asher Wolf, Journalist and Activist
Bianca Wylie, Digital Public
Tracey Gyateng, The Legal Education Foundation
Sarah Davidson, Carnegie UK
William Perrin, Carnegie UK (Trustee)

Special thank you to Sananda Sahoo, FIMS Western University, who provided valuable editorial support.

September 2022

Contents

Foreword

Pressure on public finances means that governments are trying to do more with less. Increasingly, policymakers are turning to technology to cut costs. But what if this technology doesn't work as it should?

This report looks at the rise and fall of automated decision systems (ADS). If you've tried to get medical advice over the phone recently you've got some experience of an ADS – a computer system or algorithm designed to help or replace human decision making.

These sorts of systems are being used by governments to consider when and how to act.

The stakes are high. For example, they're being used to try to detect crime and spot fraud, and to determine whether child protective services should act.

This study identifies 61 occasions across Australia, Canada, Europe, New Zealand and the United States when ADS projects were cancelled or paused.

From this evidence, we've made recommendations designed to increase transparency and to protect communities and individuals.

In addition, these suggestions are intended to enhance the innovation capacity of the public sector by recognising the necessity of strong governance and institutional review. We understand that governments are likely to develop new ADS initiatives, but this report makes the case for greater safeguards and oversight.

This work is important to Carnegie UK as we make the case for putting collective wellbeing at the centre of our decision-making.

We believe that everyone should have a voice in decisions which affect them. This research identifies many other occasions where automated decision systems have been rolled out without sufficient democratic scrutiny.

Further, serious problems involving algorithmic decision-making have eroded faith in key institutions. This report also demonstrates that there's a profound risk that these programmes could entrench existing iniquities and inequality.

As policymakers are drawn to 'smart' solutions to thorny issues, there's work required at a national and international level for greater debate and civic participation concerning their use. We want to thank all those who contributed to this important research.

Sarah Davidson
Chief Executive
Carnegie UK

Introduction

Automating Public Services: Learning from Cancelled Systems, investigates why government departments and agencies in different countries are deciding to pause or cancel their use of automated decision systems (ADS). In this report ADS are understood as technical systems designed to help or replace human decision making.[1]

Government agencies around the world are increasingly implementing ADS to aid planning, as well as to inform decisions about service delivery.[2] Government agencies are turning to these systems in attempts to be more efficient and to target resources better. The adoption of these systems is often being done in contexts where government agencies are trying to respond to public need while facing resource constraints and cuts to services. It is for this reason that previous UN Rapporteur Philip Alston on extreme poverty and human rights raised particular concerns about them being used in ways that can facilitate cuts to services.[3] When systems like this are implemented there is often little public discussion of their limits, challenges or opportunities that come with their implementation. This is despite work detailing how the use of these systems can threaten human rights as argued by UN Rapporteur Alston and the strong body of evidence demonstrating that these systems can exacerbate discrimination, inequality, wrongly limit access to services and benefits, and increase surveillance.[4] A wide range of harms have been documented and a summary of this previous work is provided in the Data Harm Record.[5]

A lack of public discussion about the uses, benefits, risks or acceptability of ADS relates in part to the fact that, in most cases, the public is not notified where and how ADS are being used. As people learn more about the kinds of harms that are occurring as a result of ADS, there are increasing calls for more transparency, accountability and efforts to ensure greater public engagement.[6] The lack of public notification is being identified as a significant problem. Numerous organisations and governmental review bodies have been calling for public registries of ADS as a basic first step, necessary to enable more public deliberation, more accountability and better oversight processes.[7]

The ability of public servants and the public more generally to make informed decisions about the use of ADS is limited when these systems are introduced without details about the potential harms and risks that can come with these systems. Research shows that the way people think and talk about the use of digital technologies, artificial intelligence and data more generally can be influenced by the dominance of voices including 'governments, think tanks, technology firms, AI investors, global management consultancies, as well as multinational corporations' without adequate representation from civil society or the public more generally.[8] As argued by David Beer, we can see the 'faith in data' all around us, which includes visions about how data analytics, such as the use of ADS, will help us address social problems more effectively and efficiently.[9] In contrast, there is too little public information available about how data systems work in practice, particularly automated decision-making support systems.

Rationale

This project report is the outcome of researching paused or cancelled government automated systems in the UK and internationally. It adds to the growing body of literature written about ADS and other technologies, emphasising the importance of acknowledging the complexity involved in technological projects and how new systems can lead to unintended consequences. The project grew out of an increasing awareness that government agencies, in different countries, were making decisions to pause or cancel their use of automated decision systems. This challenges the idea that these systems always deliver without issue. We argue that researching the factors and rationales leading to cancellation provides a means to get beyond the myths of technology to better understand its limits and acceptability. But this is not a straightforward endeavour. It is difficult to find out where and how systems are operating, let alone where the decisions not to pursue these systems are being made. However, better understanding the reasons individuals, organisations or agencies are choosing not to continue with the use of ADS is of benefit to those who are making their own decisions about whether they should implement or continue the use of such systems.

This understanding also highlights some of the precautionary steps that may be necessary for governments in making informed decisions. Our intention was not simply to explore these systems, but to provide public officials with a better understanding of the potential pitfalls when considering, planning, procuring or piloting these ADS interventions. Through this research we have identified a number of considerations for policy makers to take into account when considering, developing or implementing future ADS. We have also sought to contribute to the work being done by civil society organisations active in this space.

The motivation for this project is not to discourage technological innovation in the public sector. Instead, it is to identify why this innovation must be done in a transparent, collaborative and accountable manner and in a way that more fully recognizes consequences, both positive and negative. Better use of data can help public services run more efficiently and effectively, delivering a range of public benefits to individuals and communities and this should be recognised.[10] However, this is not easy to achieve. The increasing use of data raises concerns about privacy, security, transparency and accountability as well as the potential for discriminatory sorting, exclusion and exploitation, over-surveillance, the reinforcing of stereotypes and broader unintended consequences. Furthermore, it is clear that the pausing or cancellation of systems is not always a negative outcome. Examples of testing and learning from piloting any type of system can and should be shared with the associated learning. The challenge and part of the impetus for this work comes when these systems go live and have real-world impact without the appropriate testing or considerations in place.

This research also recognises that the way that technologies are developed and implemented are not predetermined. Changes can be made to the way these technologies are procured, designed or delivered to improve outcomes for individuals, communities and organisations. As argued by Virginia Eubanks, it is important to recognize that technologies are 'sites of struggle.'[11] To do so is to recognise that there are competing values, politics and visions informing decisions about how technologies are deployed to support public services. Paying attention to where these struggles are occurring makes it more apparent that our shared futures are not predetermined. There remains important work to be done to determine the kinds of datafied futures we want, particularly in the use of ADS.

Methods

This research was undertaken by the Data Justice Lab, with Carnegie UK providing review and support. The Data Justice Lab team included Joanna Redden, Jessica Brand, Ina Sander and Harry Warne. The Carnegie UK team included Anna Grant and Douglas White. Primary research began in February 2020[12] and concluded Sep 2020 with data analysis and report drafting occurring between October 2020 to January 2022. While publication of this final report has been delayed due to the ongoing COVID-19 pandemic, efforts have been made to ensure that the discussions and recommendations are still timely and relevant.

The first part of the project involved a scoping exercise to identify and list paused or cancelled systems in Australia, Canada, Europe, New Zealand, the United Kingdom and the United States[13] as well as to identify the range of factors influencing decisions to pause or cancel systems. In total 61 examples were identified. The research in this phase involved informal conversations and emails with experts in different countries (researchers and civil society organisations). We asked these experts to provide us with examples of paused or cancelled systems they were aware of in their own countries. We also used keyword searching to find examples that were documented in media accounts, government documents and various organisation and institutional websites. We recognise that this approach may not have given us an exhaustive list of cancelled systems, but part of the point of this research is to understand what is primarily in the public domain and easily accessible. We produced a scoping report that lists and briefly discusses why each system was paused or cancelled. These brief reports were analysed to identify the range of factors that led to cancellation as presented in research findings. A list of factors was compiled and then a content analysis of the scoping report, as well as the more detailed case study research reports (detailed below) was done to quantify our findings and identify patterns.

Another goal of the project was to conduct more detailed case study investigations that would facilitate comparison. With our case studies we wanted to be able to compare the factors and rationales leading to cancellation in similar areas of service but in different countries. Through an analysis of our scoping report we identified that we had examples of cancelled systems in different countries in the areas of fraud detection, child welfare and justice. We conducted 12 case studies in total, four in each of the policy areas identified. For each case study we collected relevant documents including legal documents, government reviews and audits, research reports and media reports. We also set out to interview two to three people with direct experience with the cancelled ADS per case study. We sent interview requests to government administrators, lawyers, politicians, civil society organisation representatives and community activists. In some cases it was challenging to find people who would agree to be interviewed. In these cases we relied heavily on published material. We interviewed 23 people in total. A case study report was written for each case study. These reports ranged from 20 to 40 pages, depending on the amount of information available. The case study reports were then condensed to summary reports (available in part two of this report) that provide a focused overview which includes: summary, why the system was stopped, key factors leading to change, how the system worked, information available about impact and information that suggests what might happen next.

Table 1: Case Studies

Fraud Detection

Netherlands	Ministry of Social Affairs and Employment System Risk Indication (SyRI) (2014-2020)
United States, Michigan	Michigan Unemployment Insurance Agency stops using Michigan Integrated Data Automated System (Midas) for automated fraud assessments (2013-2015)
Australia	Robo-debt / Online Compliance Intervention Stopped (2016 – 2019)
UK	Several local authorities stop using automated risk-based verification systems (2013 - 2020)

Policing

Germany, Baden-Württemberg	The German federal state of Baden-Württemberg stops using PRECOBS predictive policing system (2015-2019)
United States, Los Angeles	Los Angeles Police Department stops using Los Angeles Strategic Extraction and Restoration (Laser) (2011-2018) and PredPol (2009-2020)
New Zealand	The New Zealand High Tech Crime Group decide not to pursue use of Clearview facial recognition technology
UK, Durham	Durham police stop using the Harm Risk Assessment Tool (HART). (2016-2020)

Child welfare

Denmark	Denmark decides not to pursue use of the Gladsaxe model. (2017 - 2018)
United States, Illinois	Illinois Department of Children and Family Services (DCFS) stops use of Rapid Safety Feedback (RSF) (2015 - 2017)
New Zealand	Government decides not to use Predictive Risk Modelling to identify children at risk of abuse and neglect (2012 - 2015)
United Kingdom, Hackney	Hackney Council decides not to pursue use of Early Help Profiling System (2015 - 2019)

What has been cancelled? Findings summary

In this report our intention is to share insights gained both from the initial broad scoping work that looked at 61 cancelled systems and our 12 in-depth case studies. This section presents high level patterns or trends identified across the full scoping study, while the subsequent discussion section sets out deeper analysis of 12 case studies.

A short overview of each 61 examples highlighted through the scoping study is outlined in Part One: Scoping Review.

Findings from the Scoping Review

Region

As detailed in Table 2 below, our initial scoping review research identified a total of 61 ADS that were paused or cancelled. These include systems implemented by different levels of national and local government.

Table 2

Total number of cancelled systems	61
Number in Europe (excluding UK)	11
Number in the United Kingdom	10
Number in North America	33
Number in Australia and New Zealand	7

Stage of the process

We found systems were paused or cancelled at different stages (Table 3), with some systems cancelled in the investigation and development stage, some cancelled after a pilot had been done and some cancelled after the system had been implemented.

Facial recognition presented a unique range of cases because we identified a number of places where a moratorium or ban had been instituted on the use of facial recognition technologies. Many of these cases have been coded as pre-emptive actions although it is possible police forces in these places had in fact trialled facial recognition through the use of third-party systems such as Clearview AI. More research would be needed to determine with certainty that these systems were banned pre-emptively.

As indicated by Table 3, half of the systems cancelled were cancelled after they had been implemented or used. Of these 31 cases, 17 cancellations involved investigative or critical media coverage. In addition, 18 of these 31 cases were cancelled as a result of government review or legal challenge. The factors that led to cancellation operate in concert. This is discussed more below in relation to Table 5. The factors are mentioned here in relation to Table 3 to note that when a decision has been made to cancel an ADS after it has been implemented, this usually follows critical media coverage, legal interventions and review.

Table 3
Number of systems cancelled by stage

Development / Investigatory Stage	3
After Pilot / Testing	9
After Implementation / use	31
Pre-emptive ban / moratorium	18

Table 4
Systems cancelled by area of service

Justice (Policing and Law)	32
Welfare and benefits fraud detection	12
Child protection	5
Education	4
Immigration	3
Finance	2
Border control	1
Smart city	1
Health	1

Area of service

As indicated by Table 4 most of the systems we identified as being paused or cancelled are in the areas of Justice (32), in relation to welfare and benefits administration (12), in child protection (5), in education (4) and immigration (3). The high number of cases in Justice is in part due to the number of government agencies that decided to stop using facial recognition technologies (16) and the decision by Clearview AI to stop selling its facial recognition services in Canada (1). In the case of fraud detection, most of these cases relate to government agencies stopping their use of automated fraud detection systems in benefit administration.

Factors influencing cancellation

In the absence of government registries of where and how automated decision systems are being used, it is difficult to identify patterns with certainty. However, our findings support previous research[14] suggesting that government agencies are seeking to be more efficient, to cut costs and to experiment with ADS in the areas of justice and policing, welfare and benefits administration, child protection and, to a lesser extent, education and immigration.

Our research identified a range of factors, often working together, which influenced decisions to pause or cancel planned systems or systems in use.[15] For this reason, the categories posted below are not exclusive, meaning that each cancelled system could have a number of factors recorded as leading to decisions to cancel.

As detailed in Table 5, half of the decisions to pause or cancel the use of ADS relate to internal government or political concerns about the effectiveness of the systems, in other words is the system doing what it was designed to do. This finding demonstrates the need for more widespread open discussion as a matter of urgency about how the systems work in practice, given the human and financial cost when the systems do not work. Over half of the systems cancelled in relation to government concerns about effectiveness, were in the U.S. (17).

12 systems were cancelled in part as a result of critical review by governmental organisations. Five of these examples were in Canada. Here too applications in the area of justice figure prominently: Eight of the 12 cases in this area were in the area of policing and law. In almost all cases where there was a critical review by a governmental organisation there was also critical media coverage.

Table 5
Factors influencing decisions to cancel

Government agency decision – effectiveness	31
Civil society critique or protest	26
Critical media investigation	24
Legal action	19
Government concern - privacy, fairness, bias, discrimination	13
Critical government review	12
Political intervention	8
Government decision - procurement, ownership	6
Other	5
Corporate decision to cancel availability of system	3

These reviews were conducted by the Office of the Inspector General (U.S.), Attorney General (U.S.), Special Commission (U.S.), the Office of the Privacy Commissioner (CA), Senate Committee Review (CA and AU), Police Supervisory Body (BE) and Audit Office (PL). These examples are important demonstrations of the essential role that can be played by oversight bodies.

Another key finding is that nearly half of the cancelled systems had been subject to civil society critique and critical media coverage. We found that in many cases media coverage was responsible for identifying trials or implemented systems whose existence was not widely known until reported. In this way, media coverage is playing a significant role in rendering visible the systems and their impact on people. It was also common for media coverage to reference civil society concerns being raised about particular systems. Civil society critique took the form of community organisations raising concerns and research outputs that raised concerns about the impact of ADS. The key role played by civil society groups and individuals is discussed more in the following discussion section.

Nearly one third of paused or cancelled systems listed were stopped as a result of legal action (19 cases). In some cases, legal challenges were launched because it was clear that the harms or risks involved had not been sufficiently considered and mitigated against, either from deployment or in response once problems had been identified following implementation. In other cases, legal action involved city councillors introducing legislative bans on the uses of facial recognition technology. Twelve of the 19 examples of legislative action were in the U.S. Six of the 19 cases relate to applications in the area of welfare and 10 of the 19 relate to applications in the area of justice. The majority of legal challenges were brought by those negatively affected through class action lawsuits and by civil society organisations such as unions, tech justice and civil liberties groups. One legal challenge was brought by the state against IBM. One case was brought by a human rights commissioner and in another case a formal complaint was filed by an information commissioner. Nine cases involved city councils approving ordinances to ban the use of facial recognition technology or predictive policing. These bans were preceded by years of campaigning by civil society organisations. Legal actions were taken for a range of reasons: to seek damages after harms caused by errors and inaccuracies; on the basis that systems lacked fairness and due process; that they violated rights protected through the European General Data Protection Regulation; concerns about non-compliance with the European Convention of Human Rights; charges that a system was in breach of the constitution (Poland); the charge that a system was in breach of rights to protection of personal data; charges that systems were discriminatory and disproportionately negatively impacted marginalized communities.

The level of private company involvement in ADS varied across the 61 examples identified in the scoping study. Some of the ADS were developed in-house by government organisations, some were purchased and some were outsourced to third party providers. The research demonstrated that a small but important number of these systems [six] were cancelled due to governance issues including those relating to procurement and ownership.

In combination, these findings suggest there are competing understandings of acceptable data practices. Given the civil society mobilisation, legal challenges and the number of interventions by oversight bodies, our research suggests that there are competing understandings about effectiveness, impact, accountability and how data systems can infringe people's rights across areas of application. A key issue is that it takes a lot of work and resources to challenge a data system once in place.

What can we learn?

In this discussion we take the findings from the scoping review and the deeper learning from the 12 case-studies to analyse the broader implications of the work and provide recommendations to advance more meaningful civic participation, improve the structures and systems currently in place to prevent harm and to increase wellbeing of individuals and society.

Detailed discussions of the case studies are provided in Part 2 of this report. These overviews, which draw upon interviews and document analysis, provide: a summary, details about why the system was cancelled, background information, key factors identified as leading to change, information about how the system worked, details about positive or negative impacts as well as what is known about the case going forward.

Governance, oversight and accountability

Effective governance and trust in institutions is core to our democratic wellbeing and this includes the use of ADS to inform decisions. This trust is based on our collective confidence that appropriate processes are in place to ensure those decisions are made in the best interest of the public and at the same time will not cause distress or unfair outcomes for particular individuals or communities. Part of this process involves ensuring that sufficient checks and balances are in place, including the independent verification and interrogation of these systems.

Our research identifies 61 ADS that have been cancelled in various countries. This research, along with efforts to map government uses of ADS, are an indication of how quickly the use of ADS is being introduced across public sectors. Algorithm Watch have highlighted that systems of oversight are not being adapted at the same pace as these new changes.[16] The former UN Special rapporteur on extreme poverty and human rights, Philip Alston, examined the rapid uptake of digital technology in welfare state planning and administration. In his 2019 report to the UN General Assembly, he argued that his investigations into government uses of digital technology had led to concerns about the direction we, collectively, are headed. 'As humankind moves, perhaps inexorably, towards the digital welfare future it needs to alter course significantly and rapidly to avoid stumbling zombie-like into a digital welfare dystopia.'[17]

More public discussions and meaningful deliberation about the uses of ADS provide a means to work toward collective wellbeing. Such discussions depend upon the public being more informed about where and how ADS are in use. This is difficult as most government bodies, including most of those studied for this research, do not publish public registries of where and how ADS are in use. France is in the process of producing this kind of registry and registries are produced by a number of cities including Amsterdam, Helsinki and New York.[18] The UK government launched an algorithmic transparency standard in November 21, a step in the right direction. It will be important to see how the standard and the registry works in practice. Our research process makes clear the time and effort required to investigate government uses of ADS. Calls for governments around the world to establish ADS registers are widespread and have been made by a wide range of groups including the UK House of Commons Science and Technology Committee, the Ada Lovelace Institute (UK), Citizen Lab (Canada), Algorithm Watch (Germany), The Law Society (UK), the Law Commission of Ontario (Canada), the Data Justice Lab (UK), Access Now and the Open Government Partnership (OGP).[19]

Maintaining registries and archiving these would require, as suggested by Rashida Richardson et al., that government agencies dedicate sufficient resources to record and communicate about the systems clearly; make greater effort to make procurement details and company processes more transparent; explain intentions and uses of ADS more openly; and respond to citizens requests for information.[20]

Advocates of ADS registers have proposed that registers should include results of audits, details of datasets and variables being used and how the system is intended to be used. These would provide a centralised and verified space to hold ADS information, allowing not just greater oversight of how, when and why these systems are deployed but also facilitate greater learning.

Michael Veale and others, have also highlighted the importance of information and privacy commissioners being resourced to investigate algorithmic systems proactively, as opposed to operating reactively when concerns are raised (2019, p. 5).[21] Our research demonstrates the significant role oversight bodies and reviews can play where there have been concerns about negative impacts of ADS. For example, in 2020 the Office of the Privacy Commissioner of Canada announced that 'Clearview AI has advised Canadian Privacy Protection Authorities that, in response to their joint investigation, it will cease offering its facial recognition services in Canada.' The Privacy Commissioner noted that: '[t]he joint investigation was initiated in the wake of media reports which stated that Clearview AI was using its technology to collect images and make facial recognition available to law enforcement in the context of investigations.'[22] Our research provides numerous examples of oversight bodies making significant interventions through investigation: the Office of the Inspector General reviewed police use of predictive policing in LA (US); a Special Commission in MA (US) raised concerns about the use of pretrial risk assessment tool; the Attorney General in Pittsburgh banned police use of Clearview facial recognition system; Canadian privacy commissioners investigated multiple police uses of facial recognition systems and license recognition system in Vancouver; and the Belgian Supervisory Body for Police Information enquiry into airport use of facial recognition led to request for temporary pause. This research demonstrates the value of ensuring such bodies have the resources available to do these kinds of investigations as well as the ability to intervene when necessary.[23]

Recommendation

Registers of Algorithmic systems should be produced and made publicly available. These registers should be contributed to by all levels of government and maintained on a continual basis. Responsibility for the registers could lie with information officers or be set through collaborative task force as has been done in France.[24]

Recommendation

Resource public organisations, including regulators, to support efforts for greater transparency and accountability in relation to ADS.

Compounding inequalities

There is an abundance of research, including that contained in this report, which suggests that ADS (and AI systems more generally) do not always work as intended and that significant harm can be caused by them. Previous research demonstrates how ADS can be used in ways that discriminate, exacerbate inequality, infringe upon rights, socially sort, wrongly limit access to services and benefits and intensify surveillance.[25]

Furthermore, our research demonstrates that ADS are being used in the areas of justice, welfare administration, fraud detection, child protection, education and immigration. High levels of inequality already exist in how citizens experience these policy domains, making the use of technologies that could exacerbate this inequality in these contexts more dangerous.[26]

Our research findings reinforce calls for the need to be alert to how the use of ADS can create differential systems of advantage and disadvantage (Hoffman 2019, Crenshaw 1989). The potential for ADS to differentially advantage groups with more privilege while disadvantaging groups historically marginalized was a critique raised in a number of the cases identified throughout our scoping document and case study investigations. For example, concerns about how ADS compounded inequality was raised in concerns about: LA police use of PredPol and Laser; the use of facial recognition technology by authorities in Canada, the U.S., Belgium, and France; the UK Home Office use of an automated visa ranking system; the use of automated fraud detection systems in AU, U.S. and NL; and NZ plans to use predictive risk modelling in child welfare.

In understanding this research 13 American cities were identified as banning the use of facial recognition technologies because of concerns about technological bias and inaccuracy. Researchers and activists have been central in raising concerns about facial recognitions. The intersectional analysis conducted by Joy Buolamwini amd Timnit Gebru has been particularly influential. They demonstrate how these technologies can discriminate by misidentifying people of colour and women at rates much higher than white men.[27]

Misidentification through facial recognition technology has already led to wrongful arrest and detention in the United States.[28] Our case study research details how fraud detection systems have wrongly targeted thousands of people in Michigan (U.S.) and Australia leading to wrongful debt collection, persecution, bankruptcy, stress, family breakdown and illness. A fraud detection system in the Netherlands was ruled to violate human rights in 2020. All three of these fraud detection systems have been subsequently cancelled.

There is a need for those who wish to implement ADS to address the history of these kinds of applications. For example, as detailed above, our research has documented examples where the use of fraud detection systems have led to wrongful targeting and been found to violate human rights. Accounting for this history would mean that any organisation seeking to implement ADS, for example for fraud detection, would need to address how they have taken into consideration the way problems occurred when similar systems were introduced elsewhere, in this case where there are known issues of false-positive results. Our research also identified examples of government agencies deciding to cancel predictive policing and risk scoring systems in child welfare due to concerns about bias, accuracy and effectiveness. These previous experiences with this type of application could be used as a learning opportunity for others promoting similar and new applications in these areas.

There are a number of suggestions already provided about the steps that could be taken to ensure social technical histories are accounted for. Jason Lewis and colleagues stress a good way forward is to ensure systems are developed with trust, care and responsibility by engaging the relevant communities.[29] Sasha Costanza-Chock stresses the importance of collective design principles and practices.[30] Responsibility, it has been suggested by Shunryu Garvey and Harry Collins in their considerations of artificial intelligence, means that those wanting to implement technical systems need to account for history.[31] Such an accounting of history would involve going beyond stories of success told by insiders to also include criticisms, failures and unintended consequences. They argue responsibility also involves rigorous review which includes inviting critique from experts and others who hold outside positions. Collins has suggested that one way to do this is to follow the model embraced by other sciences, such as physics, and to invite the critique of experts and outsiders. Following the recommendation of Lewis this would mean inviting critique from people with experience of challenging injustice and who will be most affected by any system to be introduced.

We did find examples of agencies inviting outsider and expert critique and review in our research. The German federal state of Baden-Württemberg invited an independent external review of a predictive policing system they were trialling called PRECOBS. This review informed their decision not to implement the system. This example and the use of independent and external review should be promoted as examples of good practice, shared and learnt from.

Our findings also support conclusions about the importance of responsibility by Karen Yeung, Rapporteur for the Council of Europe's Expert Committee on human rights and automated data processing.

"If we are to take human rights seriously in a hyperconnected digital age, we cannot allow the power of our advanced digital technologies and systems, and those who develop and implement them, to be accrued and exercised without responsibility. Nations committed to protecting human rights must therefore ensure that those who wield and derive benefits from developing and deploying these technologies are held responsible for their risks and consequences. This includes obligations to ensure that there are effective and legitimate mechanisms that will operate to prevent and forestall violations to human rights which these technologies may threaten, and to attend to the health of the larger collective and shared socio-technical environment in which human rights and the rule of law are anchored."[32]

It is important to consider who is most affected and stands to be most negatively affected when things go wrong through the use of ADS. The kinds of ADS being introduced - and in some cases cancelled - are systems that sort people into different groups on the basis of scores that are produced by data analytics that draw upon a wide range of data. These social sorting systems have been criticized as being prone to bias.[33]

Some questions to be asked to enhance responsibility include:

A. In what areas of public service are ADS being implemented,

B. Who stands to be directly affected by these systems and who will not be affected,

C. Are some groups of people being affected more by the use of ADS than others,

D. Have the potential negative impacts of these systems been investigated before implementation, including potential impact(s) on the people whose data is being used and who will be most affected by system use,

E. Who has been involved in system design and decision making about whether and how ADS should be used,

F. If a system has been cancelled - at what stage was it cancelled; how long was the system kept in place after concerns were raised and what was needed to stop harmful systems.

The research highlights that there is a need for those making decisions about whether or not to use ADS to ask these types of questions about the potential impact on inequalities before embarking on the activity, rather than mitigating the effects post-intervention. A key component of this work should involve assessing the impact of these systems through more formalised impact assessments that focus on bias, inequality, rights, privacy, surveillance and that situate the use of these technologies within wider questions of context specific systemic injustice. Half of the cancelled systems identified in our report were cancelled after the systems were implemented, rather than in a controlled pilot stage. There is a 'cost' to things going wrong that includes deepening inequality, reduction of public image and trust, financial cost, as well as the strain placed on individuals and institutions across sectors when time and resources must be expended in order to redress negative effects. In response to this, a number of organisations have presented guides about how to evaluate impact. For example, see the AISP's toolkit, 'Centering Racial Equity Throughout Data Integration.'[34]

Impact assessments should not be restricted to singular applications but should consider how people can be disproportionately negatively affected by the use of ADS across public services as a total. For example, some people may be affected by multiple applications of ADS across different kinds of public services at the same time, while other people will not have their access to services or opportunities affected by the use of ADS. Since the use of ADS always comes with the risk of bias, accuracy and error, the more a person's life is mediated by the use of these kinds of systems the more there is a chance of disadvantage and harm. The use of ADS in this way creates systems of advantage for those whose lives are not affected and systems of disadvantage for those whose lives are affected by their use.[35] Concerns about data harms and datafied injustice have led to calls for government agencies to ensure legislative protections that go beyond a focus on individual rights by including protections for communities to help confront systemic injustice[36].

Recommendation

Require equalities impact assessments before the implementation of ADS interventions and recognise the need for these assessments to account for how the ADS will address issues of systemic injustice.

Recommendation

Account for ADS history. Those wishing to implement ADS need to examine how it has previously been implemented and publicly account for how previous failures are being addressed. This could be part of required impact assessments.

Responsibility for system legality, accuracy and effectiveness

Our research discusses several ADS that were cancelled after thousands of people had been wrongly identified as being overpaid benefits through algorithmic matching systems that were prone to errors. For example, this was the case with the Australian online compliance system and with the Michigan Integrated Data Automated System. In both cases, rather than the State needing to prove to those targeted that they had been overpaid benefits, it became the

responsibility of those targeted to prove they were innocent. The way ADS are being used to, in some cases, shift this burden of proof, leading to situations where citizens are labelled guilty until they prove themselves innocent, has been raised as a significant area of concern by the previous UN Special Rapporteur on Extreme Poverty and Human Rights, Philip Alston.

"The presumption of innocence is turned on its head when everyone applying for a benefit is screened for potential wrongdoing in a system of total surveillance. And in the absence of transparency about the existence and workings of automated systems, the rights to contest an adverse decision, and to seek a meaningful remedy, are illusory"
(Alston, 2019).

Another key example of concern going forward is how people with low incomes are being differentially subjected to assumptions of guilt and burdened with proving their own innocence, in the worst cases, through the use of automated fraud detection systems that are detailed in our case study investigations. While legal challenges would eventually be successful in stopping the use of a system in Australia and the United States, these cases required much time and energy and, in the process, led to much suffering.

The automated debt compliance system in Australia and the Michigan Integrated Data Automated System continued for years despite governmental reviews, internal and external scrutiny, media coverage, legal challenges and advocacy work making it clear that the systems were destroying people's lives.[37] For example, the Australian debt compliance system known as robodebt was in place for nearly four years before being stopped. In that time the automated system sent thousands of incorrect debt notices telling individuals they had received more benefits than they were entitled to and that they now needed to repay those costs. For people already facing considerable financial stress, health conditions, caring responsibilities and poor housing conditions the experience of being hounded by debt collectors and forced to pay money to the government they did not actually owe led to significant damage as people lost their housing, means of survival and suffered physically. A similar case occurred in Michigan when the government used a fraud detection system for two years despite problems with the system. National media organisations such as Time reported that people were forced into bankruptcy, lost their homes and were thrown into family crisis as a result of accusations of fraud that were made in error. Our research and media accounts suggest that the use of this ADS for fraud assessment was stopped in response to federal government pressure and a federal lawsuit.

Our case studies demonstrate the importance of, at minimum, ensuring people who are targeted by ADS have the ability to find out information about their case and access to resources to challenge decisions made.

The research also highlighted the need for burden of proof to also extend to proving effectiveness. Concerns about effectiveness and accuracy were raised in 32 of the cancelled systems identified in our report. Concerns about accuracy and effectiveness came up in relation to systems in the area of fraud detection, benefits and tax administration, justice and policing, child welfare, and education. In some cases, as for example with the Australian online compliance system and the Michigan Integrated Data Automated System, the systems continued to be used despite known issues with accuracy.

In respect to proving safety and efficacy, ADS implementation can also learn from the traditional approach to testing from over 100 years of health and safety practice in advanced democracies. The supplier of a tool (whether internally or externally developed) should be required to test it against pre-defined external characteristics to prove that it is safe for use. In the UK, a parliamentary question raised by Lord Stevenson in 2018 proved that health and safety law does apply to software.[38]

There is a need to shift the burden of proof. By this, we mean that those introducing ADS should be required to prove that they have investigated the potential harms that could occur, but also fundamentally that they are able to demonstrate the effectiveness of the changes they are implementing. This recommendation echoes similar calls by others. For example, the What Works for Children's Social Care Centre in the UK designed and tested predictive risk assessment tools for use in child welfare and

didn't find evidence that the models worked well. Their findings led them to argue that it was the responsibility of those planning or promoting the use of ADS to prove they work and to be more transparent about their systems.[39]

Recommendation

Government agencies to stop using automated systems in ways that automatically charge and infer guilt. Given the vast resources of the State, as opposed to individuals, the onus should be on the State to prove guilt and individuals should be provided with access to information about their case and resources to challenge decisions.

Recommendation

Shift the burden of proof required to implement an ADS and prove it is safe for use. Those introducing ADS should be required to demonstrate the effectiveness of the changes they are implementing.

Review the legality of uses of automated systems

Our study identified that legal action played a role in cancelling systems in roughly a third of the case studies identified. We found 10 legal challenges across the examples cited in our scoping report, eight of these were successful with one decision still pending and one system stopped before judicial review. This finding reinforces concerns being raised by civil society organisations about the legality of how ADS are deployed and calls for greater clarity and justification of lawfulness.[40] Legal challenges were successful on the basis of: damages for failed automation (Indiana); individual determination required (Michigan fugitive felon policy); lack of due process (Houston Education Assessment System); in violation of Europe's GDPR (France use of facial recognition); not compliant with European Convention of Human Rights (ND SyRI); and contravention to protection of constitutional right to privacy (Poland unemployment scoring system).

While our research demonstrates the significant role played by legal challenges, it also reveals how long it can take for change to occur through the legal process. As argued by Rashida Richardson:

"Though legal challenges to government use of ADS have been useful in shining light on the impact of these tools and mitigating some of their worst consequences, litigation is not a viable long-term solution. In addition to being a costly and slow mitigation mechanism, litigation does not always result in adequate redress to those harmed or necessary structural change in government practices and policies. Also, depending on the types of legal claims raised, liability and responsibility may not reach third-party vendors nor incentivize best practices in ADS development and design."[41]

This raises important questions about the extent to which the full range of potential legal implications of ADS are being considered prior to these systems being implemented. Recent investigative work by Robert Booth[42], for example, uncovered examples of ADS which utilises age data as part of the process, a protected characteristic under the UK Equality Act 2010. Swee Leng Harris raises a range of legal concerns about uses of automated systems in the areas of benefits administration and immigration.[43] She argues that impact assessments should also take legality into consideration. Our findings reinforce the idea that legality is an important area of concern, given the number of systems in our study that were found illegal on different grounds and in different countries. Legal challenges after systems have been implemented is a costly, slow and sub-optimal way to interrogate legality. The solution is for those implementing ADS to ensure the legality of planned systems earlier. Avoiding blind spots in this area will require involving a wide range of stakeholders, experts and citizens.

Recommendation

Those using ADS should be required to review and publicly detail how they have assessed that a proposed ADS complies with the specific legal frameworks relevant to that system.

Review procurement processes for ADS technology

Buying-in technology from outside firms can present advantages for agencies, particularly if those procuring services and systems do not have the necessary in-house expertise to develop the kinds of systems desired on their own. However, our research also demonstrates the different ways that buying in technology and services raises significant challenges.

Most notably for ADS tech, the often-proprietary nature of the technology raises significant issues. Government agencies, in some cases, are making the decision to develop partnerships with private companies to provide data about people, to analyse the data, and in some cases to also develop and co-implement the use of the automated systems. Across the cases referenced in our report there are examples of ADS that were cancelled after a trial of the system was done and in other cases after a system has been implemented. When there are public-private partnerships involved, a challenge that has been identified is that the way the system works is often claimed to be proprietary, meaning that the company providing the system owns the knowledge rights and has a commercial interest in keeping details about how the system works private to ensure profitability. When it is argued that systems are proprietary, it is difficult for people outside of government and in some cases for those inside government, to interrogate how the system works.

While individual organisations and departments cannot be expected to be expert in technology or systems design, and working with external companies can be efficient, serious questions are raised when programmes are implemented that the client and those affected by its use cannot access. For example, this was a factor influencing the Massachusetts Senate decision to not introduce the use of ADS in their bail reform efforts. A Special Commission set up to review bail reform measures noted that a proprietary system would make it difficult for defendants to challenge the results of a decision and that it would be difficult for the jurisdiction as a whole to understand decisions. The lack of transparency about how a proprietary system worked was also raised as an issue in the case of an American LA County who piloted use of a risk scoring system in the area of child welfare. The County decided not to proceed with the system because of their inability to see how variables influenced the way that families were scored.

Our research demonstrated other challenges with procurement processes such as:

- Terms of contracts - In Indiana concerns were raised about the terms of the contract which meant that even though the system did not work as promised the State was tied to the contractual arrangement. In this case thousands of people were denied food stamps, Medicaid and other benefits. The State was still expected by the company to maintain the terms of agreement. This eventually led to a lengthy and expensive legal battle where the state and company challenged each other in court;

- Inability to meet local needs - In Victoria, AU, administrators found that after costly efforts to implement an American developed automated health record system that the system did not meet local requirements. In the end the system had to be abandoned.

These challenges with procurement need to come with recognition of the context of time and resource pressures being put on public bodies to continually deliver more with less. However, this research points to the cost that can be incurred when systems do not work as intended. The costs of this clearly need to be taken into consideration and in addition more work and resources are necessary for those across the public sector to effectively interrogate plans to enter into public-private partnerships in this area.

To support them to meet these challenges, governance bodies in the public sector should review their capability to hold organisational procurement, research and development processes to account in relation to technology products. For example asking questions of their organisations such as, does anyone in the senior leadership or other governance structure have the technical knowledge or expertise required to interrogate the procurement of leading edge technology systems?

Recommendation

Review the procurement interrogation capability of government agencies in relation to ADS and ensure resourced as necessary.

Public and Civil society engagement

The research presents examples where those implementing ADS actively sought meaningful public engagement and responded to critique; as well as examples where there was little to no effort to engage civil society, affected communities or the public more generally. Our research details numerous examples where failure to engage the public about the use of ADS is raised as a significant area of concern. For example, in Boston, school officials tried to make school bus pick up times more equitable across communities by introducing an algorithmic system to help them better plan bus routes. The effort has been called admirable by the ACLU, but because public engagement was not sought from the outset the new bus routes caught parents off guard and many raised concerns about how the new schedules would lead to disruption and complication. In response to protests the new system was stopped and school officials are now engaging with parents to co-develop new schedules. Similarly, in the city of Pittsburgh concerns about racial bias led officials to cancel their use of a predictive policing system. The Pittsburgh Task Force on algorithms pointed to a lack of public engagement by those implementing the system as a key area of concern.

Our case studies present examples of community mobilization to ensure concerns about systems are addressed in cases where there has been a failure to actively engage the public. For example, in Los Angeles, the Stop LAPD Spying Coalition have been credited with: researching the impact of police uses of ADS, highlighting how PredPol and LASER disproportionately negatively affected African Americans, Latinx and people of colour; mobilizing communities to raise concerns with officials and; calling for the Office of the Inspector General to review the systems, which was done. The Office of the Inspector General report pointed to a range of problems and was one of the factors leading the LAPD to stop the use of these systems. LASER was suspended in 2019 and PredPol in 2020.

We find no consistent way that public engagement is undertaken in the cases detailed throughout the report.

Our research demonstrates a great deal of public action and concern about changing data practices. Journalists, in addition to researchers and activists, have been playing a significant role by informing the public about where and how ADS are being used as well as by investigating the impact of these systems. In nearly half of the examples we looked at we found critical media investigation as well as civil society critique. In some cases investigative reporting makes use of ADS visible which stimulates public debate, as was the case with police testing of facial recognition technology in cities in Canada, the United States and New Zealand. In other cases, public engagement is organised by members of affected communities reactively as a result of concerns after a system has been implemented, as was the case with the Stop LAPD Spying Coalition detailed above and the Not My Debt Campaign in Australia.

The cases referenced throughout this report highlight that there are different perspectives about where and how ADS should be used and that these differing perspectives exist across and within sectors. These differences demonstrate the need for engaging in genuine dialogue with communities as well as a recognition that some of this dialogue may lead to the view that there are some areas where communities decide they do not want ADS introduced.

Our findings are in line with previous research suggesting that the risk of harm when ADS go wrong in some areas of public service management and delivery are so great that the public needs to reflect on whether such systems should be implemented in these areas at all.[44] For example, Foxglove points to facial recognition as a computer aided system that is so 'intrinsically prone to harmful outcomes' that it is impossible to mitigate this harm through regulation. The cases detailed in this report demonstrate civil society mobilization in this area as we identify 13 American cities banning police use of facial recognition technology as a result of public and political concern.[45] Virginia Eubanks has demonstrated how errors in automating Medicaid can prevent people, through system errors, from receiving necessary life-saving

medication and support. In combination, legal challenges, governmental reviews as well as community mobilization demonstrate that there are competing understandings about where and how ADS should be used across public services. When there is public engagement around questions of data use, one of the options available to the public should be to decide that they do not want a particular kind of ADS used. This is already happening, as evidenced by the municipal bans on facial recognition systems.

Engaging a wide range of experts, including communities affected, before implementation is a key step that could be taken to prevent harm and rights violations. This research supports calls to account for the use of ADS in relation to the full range of human rights, which includes privacy and equality.[46]

Recommendation

Increase public engagement around ADS implementation – those seeking to trial an ADS should seek to involve the public and civil society in discussions and decisions around the use of ADS that will materially affect individuals and communities.

Recommendation

Understand the "No Go" Areas – part of the public engagement process should be to better understand the areas the public deem unacceptable for use of ADS

Politics of care

One of the narrative threads that emerges from the cases of cancelled systems outlined throughout this report is the importance of real and perceived connection or disconnection between people and institutions. This relates directly to previous work by Carnegie UK on the importance of kindness, relationships and a sense of connection to societal wellbeing.[47] As the examples of systems cancelled after implementation reveal, problems can often be traced back to a failure to recognise the differences and complexities of people and communities as well as their rights. For example, in response to harms and concerns raised about the use of an automated fraud detection system in Michigan, the use of the Michigan Integrated Data Assessment system was stopped and the organisation returned to human led assessments. This was a result of the Michigan State legislature passing a law requiring that fraud detection be done manually in 2017. While we recognise that human assessment is not without its own bias or issues, it can allow for a more discursive and relational approach to decision making if outcomes can be more transparently understood and discussed. Further work should be done to understand how the two systems can most effectively work together.

Key questions need to be asked about how an increased focus on measurement and efficiency in public administration can lead to blind spots and can reinforce deserving/undeserving concepts and us/them binaries within systems of governance. Particular attention is required with the implementation of automated and artificially intelligent systems because of the way they reinforce a distancing language through the use of categories and quantification, as well as through practices that generate automated ranking or scoring systems. Language used by these systems can also cause distress for users and present a danger that the limits of systems are overlooked because the threat appears so severe and the limits of the systems, such as inaccuracies, are not well known. For example, it has been suggested in a news report[48] that the Illinois Department of Family Services stopped use of an ADS because it was prone to error but also that the stark language the system used was alarming (e.g. predicting likelihood of "death").

Almost half of the cancelled systems identified were stopped pre-emptively as a result of concerns raised, in some cases after careful investigation by those considering implementation. This finding suggests the importance of a politics of care and taking the time to carefully consider the use of ADS before implementation. These practices are reflective of an administrative approach that prioritises societal and community wellbeing.

Our research demonstrates what a *politics of care* in datafied societies might look like. Half of the cancelled systems identified were cancelled pre-emptively as a result of concerns about bias and fairness being raised or after investigation

and a pilot was conducted. These findings suggest that good practice by government in datafied societies is tied to a politics of care, which involves *taking the time to carefully consider* the use of these systems before implementation. As an example, the German federal state of Baden-Württemberg trialled the use of predictive policing for four years before eventually deciding not to proceed with its use. Part of the investigation process included seeking an independent review and engaging a data protection officer from the outset.

Another example of good practice was being responsive to criticism. For example, the Minister of Social Development in New Zealand decided not to proceed with plans to use predictive risk modelling for child welfare after concerns were presented about its use. This example demonstrates that as government bodies pursue ADS there is a need to ensure much greater transparency throughout the process as well as enabling detailed and sober public critique early on. This kind of effort, ideally, would lead to necessary improvements, cancellation or decisions about no go areas before scandal, harm and public protest later on.

The need for greater attention to care and community has been argued by Jason Lewis as essential to preventing blind spots in computation and system development and use, and to tackling the biases that already exist.[49]

Recommendation

Ensure a politics of care approach

This involves recognition that ADS: have significant blind spots; rely upon simplified and often biased representations of people that fail to take into consideration the differences and complexities of people and communities; present distanced and often dehumanizing abstractions; intensify power imbalances and; are prone to error.

Given all of these elements, a politics care approach is one that involves: ensuring time is taken to consult and investigate if and how ADS should be used; consultation involves extended stakeholders including affected communities; provides the potential for meaningful engagement and public review which includes being responsive to criticism to emerge as well as the option to refuse use.

Recommendations

From the work undertaken through this project, we have identified 10 recommendations which we believe are necessary to improve the landscape, culture and context of ADS use in the UK at local and national level. These recommendations build on the vast amount of work that has already gone on in this space by countless committed individuals, organisations, journalists, lawyers, and academics referenced throughout this work and beyond. These recommendations have been devised with the aim of reducing the errors and issues that lead to paused or cancelled systems, but more importantly risk significantly negative outcomes for individuals. They are intended to enhance the innovation capacity of the public sector in the UK by recognising the necessity of strong governance, openness, equity, transparency, and institutional review.

- Create and maintain **Public Registries**
- Resource public organisations including regulators to support **greater transparency and accountability**
- Enhance **procurement support**
- Require **Impact Assessments** and recognise the need to **address systemic injustice**
- Review the **legality of uses** of automated systems
- Shift the **burden of proof** required to implement a ADS
- Engage the **public**
- Understand the **"No Go"** areas
- Take **responsibility in accounting for ADS history**
- Ensure a **politics of care** approach

Endnotes

1 For a descriptive list of examples see AI Now's 'Automated Decision Systems: Examples of Government Use Cases,' https://ainowinstitute.org/nycadschart.pdf
2 Increasing government uses of ADS's has been noted by the Philip Alston, the UN Special rapporteur on extreme poverty and human rights. See for example is report to the General Assembly in 2019: https://www.hrw.org/news/2019/10/17/un-protect-rights-welfare-systems-tech-overhaul#; See also OECD (2017) Embracing Innovation in Government Global Trends, February, available: https://www.oecd.org/gov/innovative-government/embracing-innovation-in-government.pdf
3 See United Nations Human Rights, Office of the High Commissioner Release: https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25156
4 See for example work by: Gandy (2005); Lum and Isaac (2016); Lewis et al. 2020; Lyon 2002; Bennett et al. 2015; Gangardharan et al. (2014); Eubanks (2018); O'Neil (2016); Keddell (2015); Gillingham (2016); Yeung 2019; Metcalf and Dencik (2019); Redden et al. 2020; Hu (2016).
5 Two of the authors of this report have been involved in recording published accounts of data harms. The record can be found here: https://datajusticelab.org/data-harm-record/
6 For a list of the range of harms being identified see the Data Harm Record.
7 Some government agencies have started to introduce public registries. For example: Helsinki, Amsterdam, New York City, as well as some government bodies in France
8 See Clea D. Bourne, (2019) 'AI cheerleaders: Public relations, neoliberalism and artificial intelligence. Public Relations Inquiry,' 8(2), pp. 109-125.
9 David Beer (2019) The Data Gaze: Capitalism, Power and Perception, London: Sage.
10 https://www.carnegieuktrust.org.uk/publications/data-for-public-benefit/
11 Virginia Eubanks (2011) Digital Dead End. Cambridge: MIT Press.
12 All data correct as of September 2020.
13 These regions were selected to allow for comparable analysis and to be able to utilise the lessons learnt to inform UK decision making. The regions were also selected on the basis of languages that could be read and spoken by the research team.
14 Dencik, L., Hintz, A., Redden, J., and Warne, H. (2018) Data Scores as Governance: Investigating Uses of Citizen Scoring in Public Services, Data Justice Lab, https://datajusticelab.org/wp-content/uploads/2018/12/data-scores-as-governance-project-report2.pdf; Eubanks, Virginia. (2018). Automating inequality. New York, NY: Macmillan; O'Neill, Cathy. (2016) Weapons of math destruction. New York, NY: Crown.
15 For our scoping exercise, we recorded factors through our desk research. This presents limitations as it is difficult to determine with absolute certainty the full range of factors influencing decisions without detailed case study investigations. In some cases the influence of factors can be clear, for example we identified in a number of cases that there were legal orders for systems to be stopped. More difficult, particularly when the case is not recent, is determining the influence of media coverage on decision-making as drawing a causal link in such cases is not as straightforward. In some cases it was clear that media coverage led to greater awareness that a system was in use or being trialled, and there then followed an intervention from authorities to pause a system until further investigation. In other cases media coverage is linked to a broader range of factors that appear to work in concert. Our case studies, detailed in the following section, provided us with an opportunity to do more in depth investigations.
16 Loi, M., Mätzener, A., Müller, A. and Spielkamp, M. (2021) Automated Decision-Making Systems in the Public Sector: An Impact Assessment Tool for Public Authorities, Algorithm Watch, https://algorithmwatch.org/en/wp-content/uploads/2021/06/ADS-in-the-Public-Sector-Impact-Assessment-Tool-AlgorithmWatch-June-2021.pdf
17 United Nations Human Rights, Office of the High Commissioner, Announcement of Alston's report to the General Assembly in 2019. Announcement and Report can be accessed here: https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25156
18 See for example: https://ai.hel.fi/en/ai-register/; https://algoritmeregister.amsterdam.nl/; https://www1.nyc.gov/assets/ampo/downloads/pdf/AMPO-CY-2020-Agency-Compliance-Reporting.pdf; https://guides.etalab.gouv.fr/algorithmes/inventaire/#ressources-panorama-d-inventaires-existants
19 Science and Technology Committee call https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/351/35108.htm#_idTextAnchor087; Ada Lovelace Institute call https://www.adalovelaceinstitute.org/event/mandatory-reporting-public-sector-algorithmic-accountability/; Citizen Lab call https://citizenlab.ca/wp-content/uploads/2020/09/To-Surveil-and-Predict.pdf; Algorithm Watch call https://algorithmwatch.org/en/wp-content/uploads/2021/06/ADS-in-the-Public-Sector-Impact-Assessment-Tool-AlgorithmWatch-June-2021.pdf; Law Society (UK) call https://michae.lv/static/papers/2019algorithmsjusticesystem.pdf; Law Commission of Ontario call https://www.lco-cdo.org/wp-content/uploads/2021/04/LCO-Regulating-AI-Critical-Issues-and-Choices-Toronto-April-2021-1.pdf; Data Justice Lab call https://datajustice.files.wordpress.com/2018/12/data-scores-as-governance-project-report2.pdf; Access Now call https://www.accessnow.org/trust-and-excellence-the-eu-is-missing-the-mark-again-on-ai-and-human-rights/; AI Now, Open Government Partnership and Ada Lovelace Institute https://www.opengovpartnership.org/wp-content/uploads/2021/08/algorithmic-accountability-public-sector.pdf
20 The detailed list of recommendations can be found here: Rashida Richardson, ed., "Confronting Black Boxes: A Shadow Report of the New York City Automated Decision System Task Force," AI Now Institute, December 4, 2019, https:// ainowinstitute.org/ads-shadowreport-2019.html.
21 M. Veale et al. (2019) Algorithms in the Criminal Justice System, Report by the Law Society Commission on the Use of Algorithms in the Justice System, The Law Society of England and Wales, June, available file:///Volumes/NO%20NAME/Removable%20Disk/Carnegie/algorithms-in-criminal-justice-system-report-2019.pdf
22 For details see: https://www.priv.gc.ca/en/opc-news/news-and-announcements/2020/nr-c_200706/; https://

www.priv.gc.ca/en/opc-news/news-and-announcements/2020/an_200221/

23 This has been suggested by many, most recently by the Law Society Commission review algorithms in the justice system, a coalition of civil society groups in New York in their shadow report of the city's automated decision system task force

24 See: https://www.opengovpartnership.org/stories/building-public-algorithm-registers-lessons-learned-from-the-french-approach/

25 See: Benjamin, R. (2019). Race After Technology: Abolitionist Tools for the New Jim Code. Cambridge: Polity Press; Eubanks, V. (2018). Automating Inequality. New York: Macmillan; Gandy, O. (1993). The Panoptic Sort: A Political Economy of Personal Information. New York: Harper Collins; Lyon, D. (2002). Surveillance as Social Sorting: Privacy, Risk and Automated Discrimination. New York, NY: Routledge; Lum, K. and Isaac, W. (2016). To Predict and Serve. Significance. 13(5): 14-19. Retrieved from https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00960.x; Gangadharan, S. P., Eubanks, V., and Barocas, S. (2014). Data and discrimination: collected essays. Open Technology Institute and New America. Retrieved from https://www.ftc.gov/system/files/documents/public_comments/2014/10/00078-92938.pdf (9 Sept. 2015); Hu, M. (2015). Big Data Blacklisting. Florida Law Review, 67: 1735-1809; O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York, NY: Crown Publishing; Keddell, E. (2015). The Ethics of Predictive Risk Modelling in the Aotearoa/New Zealand Child Welfare Context: Child Abuse Prevention or neo-Liberal Tool? Critical Social Policy, 35 (1): 69–88. doi:10.1177/0261018314543224; Stark, L. (2018). Algorithmic Psychometrics and the Scalable Subject. Social Studies of Science. 48(2), 204–231; Lewis, J. E., ed. (2020). Indigenous Protocol and Artificial Intelligence [Position Paper]. Honolulu, Hawaii: The Initiative for Indigenous Futures and the Canadian Institute for Advanced Research (CIFAR).

26 Research summaries of inequality, systemic and structural racism across these areas include: U.S. https://www.sentencingproject.org/publications/un-report-on-racial-disparities/; https://www.ohchr.org/EN/NewsEvents/Pages/systemic_racism.aspx, CA https://www.canada.ca/en/department-justice/news/2021/06/addressing-systemic-racism-against-indigenous-peoples-in-the-justice-system-government-of-canada-investments-and-initiatives.html; AU https://www.alrc.gov.au/publication/pathways-to-justice-inquiry-into-the-incarceration-rate-of-aboriginal-and-torres-strait-islander-peoples-alrc-report-133/;https://www.sbs.com.au/nitv/nitv-news/article/2018/03/14/institutional-racism-evident-australian-justice-system; UK https://warwick.ac.uk/services/sg/si/diversity/advance_he_-_understanding_racism_report.pdf; EU https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/690525/EPRS_BRI(2021)690525_EN.pdf; NZ https://www.hrc.co.nz/news/migrant-new-zealanders-experience-fear-loss-culture-and-identity-due-racism/

27 Work by Joy Bulamwini and Timnit Gebru (2018) has been central to making this clear: http://gendershades.org/verview.html

28 See: https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html

29 Lewis, Jason Edward, ed. 2020. Indigenous Protocol and Artificial Intelligence Position Paper. Honolulu, Hawai'i: The Initiative for Indigenous Futures and the Canadian Institute for Advanced Research (CIFAR), https://www.indigenous-ai.net/position-paper

30 Sasha Costanza-Chock (2020) Design Justice: Community Led Practices to Build the Worlds We Need, Cambridge MA: MIT Press. Available online open access: https://design-justice.pubpub.org/

31 Jason Edward Lewis, ed. 2020. Indigenous Protocol and Artificial Intelligence Position Paper. Honolulu, Hawai'i: The Initiative for Indigenous Futures and the Canadian Institute for Advanced Research (CIFAR), https://spectrum.library.concordia.ca/986506/7/Indigenous_Protocol_and_AI_2020.pdf; Shunryu Colin Garvey (2021) Unsavory medicine for technological civilization: Introducing 'Artificial Intelligence & its Discontents', Interdisciplinary Science Reviews, 46:1-2, 1-18, DOI: 10.1080/03080188.2020.1840820; Harry Collins (2021) The science of artificial intelligence and its critics, Interdisciplinary Science Reviews, 46:1-2, 53-70, DOI: 10.1080/03080188.2020.1840821

32 Yeung, Karen, A Study of the Implications of Advanced Digital Technologies (Including AI Systems) for the Concept of Responsibility Within a Human Rights Framework (November 9, 2018). MSI-AUT (2018) 05 , Available at SSRN: https://ssrn.com/abstract=3286027

33 Gandy, Oscar. 2005. "Data Mining, Surveillance, and Discrimination in the Post-9/11 Environment." In The New Politics of Surveillance and Visibility, edited by K. D. Haggerty, and R. V. Ericson, 363–384. Toronto: University of Toronto Press.; Lyon, David, ed. 2002. Surveillance as Social Sorting: Privacy, Risk and Automated Discrimination. New York, NY: Routledge; Anna Lauren Hoffmann (2019) Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse, Information, Communication & Society, 22:7, 900-915, DOI: 10.1080/1369118X.2019.1573912

34 AISP (2021) Centering Racial Equity, https://www.aisp.upenn.edu/centering-equity/

35 For discussion on this see: Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. University of Chicago Legal Forum, 1(8), 139–167; Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against Women of Color. Stanford Law Review, 43(6), 1241–1299; Hoffman, A.L. (2019) Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse, Information, Communication & Society, 22:7, 900-915, DOI: 10.1080/1369118X.2019.1573912

36 For example, Jennifer Raso has argued for changes to administrative law in order to address systemic injustice: Raso, Jennifer, AI and Administrative Law (November 20, 2020). in Florian Martin-Bariteau & Teresa Scassa, eds., Artificial Intelligence and the Law in Canada (Toronto: LexisNexis Canada, 2021), Available at SSRN: https://ssrn.com/abstract=3734656 or http://dx.doi.org/10.2139/ssrn.3734656

37 There is a great deal of media coverage documenting the hardship experienced by those wrongfully targeted by the Australian and Michigan system. For example, see this article by The Guardian's Henrique-Gomes https://www.theguardian.com/australia-news/2021/jun/11/robodebt-court-approves-18bn-settlement-for-victims-of-governments-shameful-failure as well as this Time magazine article by De La Garza https://time.com/5840609/algorithm-unemployment/

38 See https://questions-statements.parliament.uk/written-questions/detail/2018-05-23/HL8200

39 See Clayton, V., Sanders, M., Schoenwald, E., Surkis, L., Gibbons, D. (2020) Machine Learning in Children's Services Summary Report. What Works for Children's Social Care, September, available at: https://whatworks-csc.org.uk/

wp-content/uploads/WWCSC_machine_learning_in_childrens_services_does_it_work_Sep_2020_Accessible.pdf; Sanders, M. (2020) Machine Learning; Now is a Time to Stop and Think, What Works for Children's Social Care. 10 September. Available at: https://whatworks-csc.org.uk/blog/machine-learning-now-is-a-time-to-stop-and-think/

40 See for example M. Veale et al. (2019) Algorithms in the Criminal Justice System, Report by the Law Society Commission on the Use of Algorithms in the Justice System, The Law Society of England and Wales, June, available file:///Volumes/NO%20NAME/Removable%20Disk/Carnegie/algorithms-in-criminal-justice-system-report-2019.pdf

41 Rashida Richardson (2019) Written Testimony for New York City Council Committee on Technology, January 22, AI Now Institute, p. 4, available: https://ainowinstitute.org/ainow-rashida-richardson-testimony-nycc-ads-disclosures.pdf

42 Robert Booth (2021) Calls for Legal Review of UK Welfare Screening System Which Factors in Age, The Guardian, 18 July, available: https://www.theguardian.com/society/2021/jul/18/calls-for-legal-review-of-uk-welfare-screening-system-that-factors-in-age

43 Swee Leng Harris (2020). Data Protection Impact Assessments as rule of law governance mechanisms. Data & Policy, 2, E2. doi:10.1017/dap.2020.3

44 This is discussed in AI Now's 2018 report: D. Reisman, J. Schultz, K. Crawford and M. Whittaker, Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability, available https://ainowinstitute.org/aiareport2018.pdf. It is also raised by Virginia Eubanks in Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor, New York: St. Martin's Press.

45 Foxglove, Data in government: making it fair and legal (March 2020) Available: https://www.foxglove.org.uk/news/data-in-government-making-it-fair-and-legal

46 Recommendations that relate to specific calls for human rights implications to be addressed can be found here: Bailey, J., Burkell, J., Steeves, V. (2020) "AI technologies -- like police facial recognition - discriminate against people of colour," The Conversation, August, https://theconversation.com/ai-technologies-like-police-facial-recognition-discriminate-against-people-of-colour-143227; The State of Data 2020 report: mapping children's data in state education in England (2020) defenddigitalme https://defenddigitalme.org/the-state-of-data-2020/ part 1; Petra Molnar and Lex Gill (2018) "Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System" International Human Rights Program (Faculty of Law, University of Toronto) and the Citizen Lab (Munk School of Global Affairs and Public Policy, University of Toronto); Robertson, K., Khoo, C., Song, Y. (2020) "To Surveil and Predict: A Human Rights Analysis of Algorithmic Policing in Canada," Citizen Lab (Munk School of Global Affairs & Public Policy, University of Toronto) and the International Human Rights Program (Faculty of Law, University of Toronto), https://citizenlab.ca/wp-content/uploads/2020/09/To-Surveil-and-Predict.pdf https://citizenlab.ca/wp-content/uploads/2018/09/IHRP-Automated-Systems-Report-Web-V2.pdf

47 See Carnegie UK, Kindness Brief, https://www.carnegieuktrust.org.uk/project/kinder-communities/. In particular the work by Julia Unwin (2018) Kindness, emotions and human agency: The blind spot in public policy, Carnegie UK, https://www.carnegieuktrust.org.uk/publications/kindness-emotions-and-human-relationships-the-blind-spot-in-public-policy/

48 See David Jackson and Gary Marx (2017) Data mining program designed to predict child abuse proves unreliable, DCFS says, Chicago Tribune, available: https://www.chicagotribune.com/investigations/ct-dcfs-eckerd-met-20171206-story.html

49 Lews, J.E. (2021) "From Impoverished Intelligence to Abundant Intelligences," Medium, May, available: https://jasonedwardlewis.medium.com/from-impoverished-intelligence-to-abundant-intelligences-90559f718e7f

50 Professor Mukherjee is Director of Columbia Law School's Immigrants' Rights Clinic. She was interviewed by The Atlantic following her visit to a detention centre in 2019. You can read her interview here: https://www.theatlantic.com/family/archive/2019/06/child-detention-centers-immigration-attorney-interview/592540/

51 See statement by PILP here: https://pilpnjcm.nl/en/dossiers/profiling-and-syri/

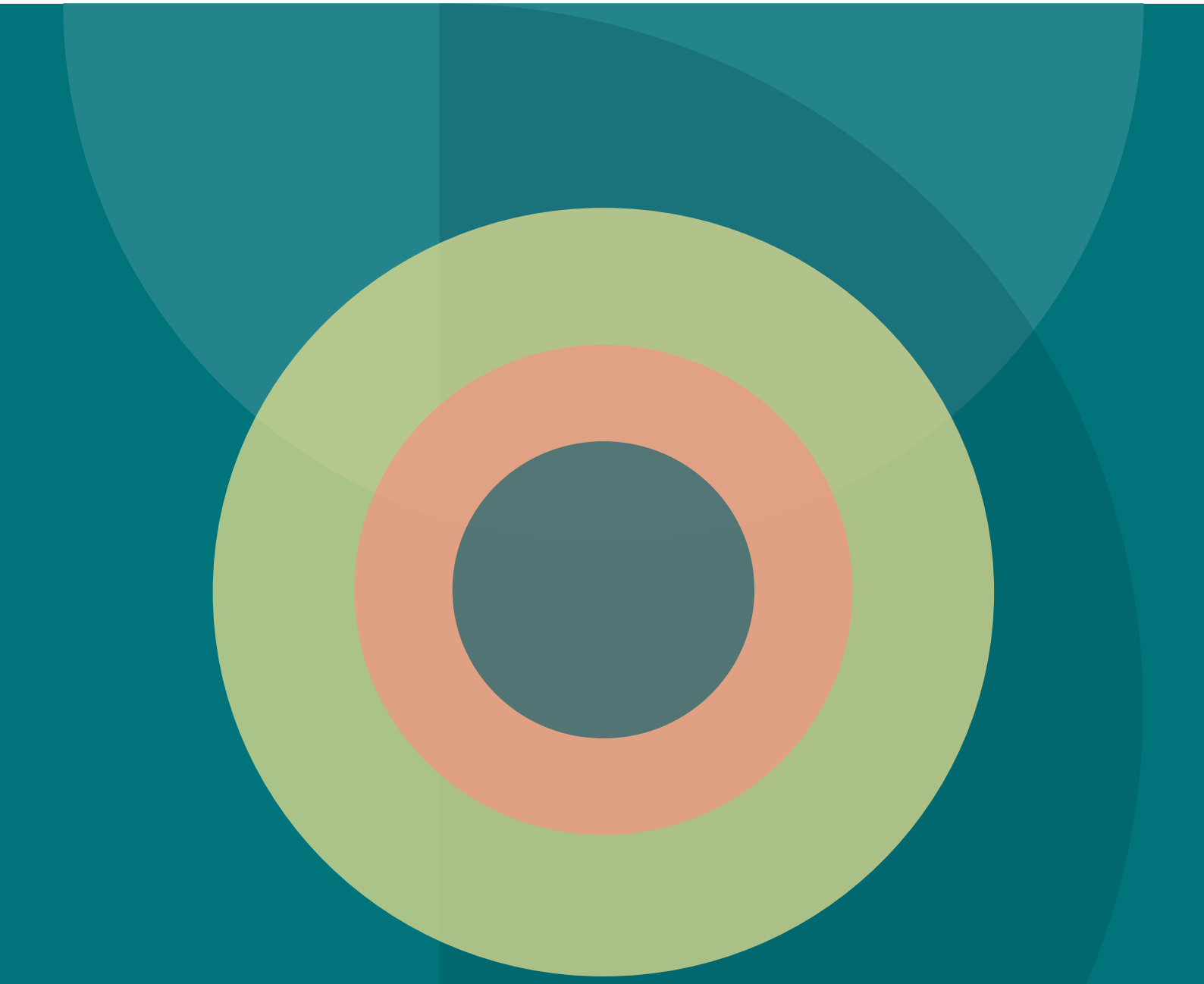
52 Ilja Braun (2018) 'High Risk Citizens,' Algorithm Watch, 4 July 2018, available: https://algorithmwatch.org/en/high-risk-citizens/

53 van Dalen, Steven; Gilder, Alexander; Hooydonk, Eric; Ponsen, Marc (2016) System Risk Indication: An Assessment of hte Dutch Anti-Fraud System in the Context of Data Protection and Profiling. Public Interest Litigation Project, available: https://pilpnjcm.nl/wp-content/uploads/2016/06/memorandum_1_-_system_risk_indication_an_assessment_of_the_dutch_anti-fraud_system_in_the_context_of_data_protection_and_profiling-1.pdf

collective wellbeing Carnegie UK

Data Justice Lab

Western FIMS



Automating Public Services: Learning from Cancelled Systems Scoping Report

Joanna Redden, Jessica Brand, Ina Sander and Harry Warne
Data Justice Lab

Contents

Introduction

This section provides the results of our scoping work. We set out to find out how many government agencies in Australia, Canada, Europe, New Zealand and the United States have decided to pause or cancel their use of automated decision aid systems.
In this scoping section we provide brief summaries of paused or cancelled systems with links to additional resources readers can access if they want to pursue more detailed information. A more comprehensive exploration of examples are provided in the next section of this report, which provides case study overviews.

USA

Indiana, IBM contract for the automation of Indiana's welfare services cancelled

Summary

In 2006, the state of Indiana's Family and Social Services Administration signed a 10-year contract with IBM and the call center company Affiliated Computer Services (ACS), to automate its public assistance eligibility processes. In October 2009 Indiana's Governor Mitch Daniels then cancelled the $1.3 billion contract to privatise and automate the processing of applications for food stamps, Medicaid and other benefits received by more than a million Indiana children, seniors, elderly and disabled residents.

Why was it dropped?

The contract was pulled in late 2009, less than three years into the 10-year deal, following complaints about long wait times, lost documents and improper rejections.

Further notes

To meet contract demands for faster eligibility decisions, IBM's data management system and call center workers denied thousands of recipients assistance. Indiana sought millions of dollars in damages from IBM, in a case that was eventually ruled in the state's favor by the Indiana Supreme Court.

Sources:

Virginia Eubanks: Automating Inequality, research published in NPR and The Nation
Michigan Department of Health and Human Services matching algorithm in "fugitive felon" police policy

Michigan Department of Health and Human Services matching algorithm in "fugitive felon" police policy

Summary

Michigan Department of Health and Human Services (MDHHS) used a matching algorithm to disqualify those with outstanding felony warrants from accessing food assistance. "Between December 2012 and January 2015, the new algorithmic system improperly matched more than 19,000 Michigan residents, and automatically disqualified each of them from food assistance benefits with this notice: 'You or a member of your group is not eligible for assistance due to a criminal justice disqualification...Please contact your local law enforcement to resolve.'" (See AI Now report, p.19.)

Why was it dropped?

As a result of a class-action lawsuit filed on behalf of those receiving disqualification notices. The Sixth Circuit upheld a federal district court ruling "enjoining the State's inadequate notices and any disqualification based on computer matching without an individualized determination" (See AI Now report, p. 19).

Sources

AI Now Institute - Litigating Algorithms: Challenging Government Use of Algorithmic Decision Systems

Michigan Unemployment Insurance Agency stops using Michigan Integrated Data Automated System (Midas) for automated fraud assessments

(Case study discussion in following section)

Summary:

In 2013 Michigan's Unemployment Insurance Agency (UIA) launched MiDAS to administer and process unemployment benefits. It has been reported that the number of people suspected of fraud grew drastically after the system was implemented and that within two years the MiDAS system generated false accusations of fraud for 40,000 people. There have been numerous court proceedings related to people's experiences.

Why was it dropped?

It has been argued that UIA stopped using MiDAS for automated fraud assessment in September 2015 due to federal government pressure and a federal lawsuit (Charette 2018; Fleming and Fournier 2015). The UIA apologised in January 2017.

Sources:

The American Federation of State, County and Municipal Employees web archive.
Research: IEEE

Documents: Legal document and Audit report

Houston Independent School District (HISD) forced to abandon SAS's Educational Value-Added Assessment System

Summary:

In 2010, the Houston Independent School District (HISD) began using SAS's Educational Value-Added Assessment System (EVAAS). The system promised to improve teaching quality by providing standardized assessments of teachers. The district ended its contract with SAS in 2016 and did not replace the teacher-evaluation system with other software.

Why was it dropped?

A lawsuit was filed against HISD by teachers and union members. ruled in favour of the plaintiffs on procedural due process grounds.

Further notes:

In case records it is indicated that the system was going to be used to fire teachers deemed ineffective. The Houston Federal of Teachers Local 2415 and six ISD teachers argued that these systems violate 14th amendment rights to due process. SAS said its algorithm was a trade secret, which meant teachers couldn't see the evidence being used to fire them.

Sources:

Academic paper from a Professor of Educational Policy which discusses the federal lawsuit Houston Federation of Teachers v. Houston Independent School District.

Research: AI Now Litigating Algorithms
Media: Houston Chronicle
American Federation of Teachers
Legal: Judgement, Settlement Agreement

LA County drops pilot Project AURA (Approach to Understanding Risk Assessment)

Summary:

In 2014 LA County's Department of Children and Family Services worked with tech vendor SAS to "develop a pilot project that applied advanced analytic methods to data to generate a score that identified the likelihood of a tragic outcome occurring for children in contact with the Department" (Nash 2017). An LA County Office of Child Protection report notes that AURA was not used on actual cases, but tested by making use of past reports alleging child abuse or neglect. The pilot concluded in October 2014.

Why was it dropped?

According to an NCCPR blog post there was alarm at the 95 percent false positive rate. In his report to the board members of the Office of Child Protection, Michael Nash states that such a high false positive rate is problematic "because of its potential for overwhelming a system that cannot respond to such a high number of false positives and still be effective." Nash also cites as problematic the fact that the SAS system was proprietary, saying "it was a closed 'black box' model lacking transparency about how variables influenced scores."

However the report also recommends that DCFS "should continue its ongoing efforts to explore the use of predictive analytics." It also looks like the Children's Data Network and California's Department for Social Services (CDSS) have partnered with the Office for Child Abuse and Protection to research the proof-of-concept of predictive modelling in child welfare, with a view to potentially developing an in-house solution, off the back of the learning outcomes of the failed SAS pilot.

Sources

National Coalition for Child Protection Reform blog
County of Los Angeles Office of Child Protection report

Illinois Department of Family Services abandoned predictive analytics programme (Case study discussion in following section)

Summary:

In 2015 Illinois Department of Children and Family Services (DCFS) implemented Rapid Safety Feedback (RSF), a predictive analytics tool developed by the non-profit Eckerd Connects and its for-profit partner Mindshare Technology. It was brought in by former DCFS Director, George Sheldon, who was hired to address Illinois' increase in child deaths. The implementation of RSF was central to his reform plans.

Why was it dropped?

Unreliability, inaccuracy and overprediction which overloaded caseworkers with new cases; issues with the contract including non-transparent bidding; the stark language the system uses (e.g. predicting likelihood of "death") was also cited as alarming for child welfare agencies. It was dropped in December 2017.

Sources:

Media articles: Governing and Chicago Tribune

LA Police Department stops using LASER and PredPol predictive policing programme

(Case study discussion in following section)

Summary:

The Los Angeles Police Department (LAPD) began using Los Angeles Strategic Extraction and Restoration, or Operation LASER, in 2011. It was developed by Palantir. LAPD has pioneered the use of predictive policing and also started using PredPol in 2009. is an algorithm developed by the LAPD in collaboration with local universities.

Why was it dropped?

In 2018 the Stop LA Police Dept Spying Coalition made a demand in its critical report of predictive policing "Before the bullet hits the body" that the Office of the Inspector General (OIG) should conduct a review of data driven policing strategies used by the LAPD. This was heeded by the OIG and in March 2019 the report from Inspector General Mark Smith's internal audit pointed to a number of problems including lack of oversight and inconsistent criteria used to predict crime. The audit also raised concerns about how suspects were racially identified.

According to Muckrock, LASER was suspended in August 2018. But this was not discovered until Spring 2019 as the LAPD did not publicly share this information until then. The LA Times reports that "the move came after a meeting Tuesday [9th April 2019] at which members of the department's civilian oversight panel questioned the effectiveness of data-driven strategies." The LA Times also writes that Josh Rubenstein, the LAPD's chief spokesman, said "We discontinued LASER because we want to reassess the data. It was inconsistent. We're pulling back."

In April 2020, the LAPD announced that they would stop using PredPol. It was reported in BuzzFeed that Police Chief Michael R. Moore said the police would stop using PredPol due to COVID-19 related financial constraints. Campaign coordinator for the Stop LAPD Spying Coalition said he thinks the group's organizing prompted the LAPD to stop using PredPol.

Sources

Audit report from Inspector General March 2019
Report on LA predictive policing from the Stop LA Police Department Spying Coalition, 2018
This is a detailed article which gives more of an indication as to why LASER was scrapped and PredPol was not.
UCLA letter to LAPD Chief Michael Moore concerning PredPol.

Spokane Washington abandons SAFER pretrial risk assessment tool in favour of simpler, off-the shelf tool PSA

Summary:
In 2017 Spokane city and county court systems started using a new risk assessment tool called the Spokane Assessment for Evaluation and Risk (SAFER) developed by Washington State University. The system was designed to free up space in jail by making sure people were not held there simply because they are too poor to pay bail. The tool was brought in as part of a package of criminal justice reforms to reduce jail overcrowding and eliminate racial disparities in the justice system.

Why was it dropped?
Because of staff turnover, software glitches and the challenge of syncing the tool with state court data, the programme never worked as intended. After nearly three years of testing and tinkering, officials scrapped the SAFER tool in favour of a simpler, off-the-shelf programme called the Public Safety Assessment, or PSA, which requires less raw information.

Further notes:
PSA was developed by the Laura and John Arnold Foundation and is currently in use statewide in Kentucky, Arizona, New Jersey and Utah, and in counties in at least a dozen other states.

Sources
Media - Muckrock, The Spokesman

Boston School District/Boston Public Schools decides not to implement MIT algorithm for optimising school bus routes and times

Summary:
In 2016 Boston Public Schools launched a national competition to see who could come up with an algorithm that would help them improve bus routes to increase efficiency (Bertsimas et al. 2020). The winning MIT team developed an algorithm to do that and then also developed an algorithm to try and make bussing and school start times more equitable across communities.

Why was it dropped?
The solution that was put forward would have led to big changes in school start times. Parents protested and argued that such significant changes to start times would be a major disruption to their family lives, present challenges and affect their ability to work. In response to the protest the mayor cancelled the program.

Further notes:
Ellen P. Goodman has argued that the science in this case is good and the intentions admirable, she argues that this case points to the importance of public engagement.

Sources
Media - The Boston Globe
Research - INFORMS, The Regulatory Review

Chef software CEO decides not to renew contract with US Immigration and Customs Enforcement (ICE) due to employee and customer dissent amid #NoTechForICE movement

Summary

There has been ongoing protest in response to former President Donald Trump's family separation policy. The policy meant children were separated from their parents who are seeking asylum. Children and parents are held separately at different detention centres in conditions that have been widely criticized. For example, Professor Elora Mukherjee who visited one of these centres has described the conditions as appalling, degrading and inhumane.[1] In response to this there is mounting pressure being placed on companies to not work with US Immigrant and Customs Enforcement (ICE) and Customs Border Protection (CBP). One example is the group mobilizing around #NoTechForICE.
In September 2019 the CEO of the Seattle based software vendor Chef responded to internal and external pressure to stop working with ICE by announcing that the company would be cutting their ties with ICE and not renewing their contract with them.

Why was it dropped?

CEO Barry Crist wrote to his employees on Chef's website blog: "After deep introspection and dialog within Chef, we will not renew our current contracts with ICE and CBP [Customs and Border Protection] when they expire over the next year. Chef will fulfill our full obligations under the current contracts.
In addition, Chef "will donate an amount equivalent to our 2019 revenues from these two contracts directed to charities that help vulnerable people impacted by the policy of family separation and detention."

Further notes of interest:

According to Mijente, Chef is the first tech company to publicly sever ties with the U.S. agencies in response to pressure from activists and employees.

Sources

Company: Chef blog

Media: Wired

1 Professor Mukherjee is Director of Columbia Law School's Immigrants' Rights Clinic. She was interviewed by The Atlantic following her visit to a detention centre in 2019. You can read her interview here.

San Francisco, Berkeley, Alameda and Oakland CA; Boston, Brookline, Northampton, Somerville, Springfield and Cambridge MA; Portland, Ore; Portland, Maine ban use of facial recognition technologies by police departments and other public agencies

Summary

In response to widespread concerns about how facial recognition technologies are discriminatory and present a range of human rights concerns a number of cities in California, Maine, Massachusetts and Oregon have banned the use of these systems. It has been reported that San Francisco was the first American city to ban use of the tool.

Why was it dropped?

A city supervisor who supported the ban is quoted as saying that as a technological centre for the country: "We have an outsize responsibility to regulate the excesses of technology precisely because they are headquartered here."

Researchers and civil society organisations raise concerns about privacy laws not being up to date enough to ensure adequate protection of citizen rights. Also, that the inconsistencies and inaccuracies of the technology put people at risk, particularly those who are Black, Indigenous or people of colour. The ACLU has argued that facial recognition technology poses a threat to privacy, free speech and racial and gender justice.

In mid-August 2019 the ACLU "released test results showing that facial recognition software incorrectly "matched" 26 California state lawmakers with photos from a database of arrest photos, echoing research referenced above by Joy Buolamwini and Timnit Gebru showing inaccuracy on the basis of gender and skin type, with the most misclassified group being darker skinned females (Buolamwini and Gebru 2018). Lindsey Barrett notes that facial recognition systems are also less accurate for non-binary and transgender people, children and the elderly (Barrett 2020). The ACLU also raises concerns about the ability of facial recognition technologies to be used to track populations.

One of the California Assembly members to have his face falsely matched, Phil Ting, introduced a bill to block law enforcement from using facial recognition on body cameras for three years. The bill received bi-partisan support and became law in October 2019.

Similarly, in Massachusetts, Sen. Cynthia Stone Creem (D-Newton), has submitted a Bill that if passed would ban government uses of facial recognition technologies. The Bill has been referred to the Joint Committee of the Judiciary as of Nov. 2020. It is arguable that debate surrounding this legal intervention as well as related civil society mobilizing around the topic in the State contributed to cities across Massachusetts banning the uses of the technology.

It can be argued that the ban on government uses of facial recognition more broadly is linked to this earlier mobilization.

Sources:

Media: New York Times, BBC, Vox.com, Harvard Crimson, NPR, Vice, East Bay Times
Civil Society: ACLU, EFF

Massachusetts rejects recommendation of pretrial risk assessment tool in bail reform report

Summary:

In 2017 Massachusetts Senate recommended the introduction of risk assessment tools in the pretrial stage of criminal proceedings. Concerns about these tools have been made public across the United States in recent years. A Special Commission was set up to evaluate a number of changes being proposed for the bail system, including the use of risk assessment tools. In December 2019, the committee argued that concerns about bias and effectiveness influenced their decision not to implement risk assessment tool.

Why was it rejected?

The report lists six distinct reasons for its decision to reject the recommendation to use a risk assessment tool. Massachusetts has a separate statute governing the assessment of potential threats to the community, so any risk tools that predict future criminal activity in a bail determination could not be used - "only a tool used to predict the likelihood of a missed court data would be appropriate" (p.8 of report). Secondly, the report states that "risk assessment tools may have their own limitations due to their reliance on data of questionable correlation to predictability and the rigidity of application that restricts a judge's discretion." Thirdly, the report points out that because Massachusetts has a low failure-to-appear at court rate the risk assessment tool is less useful and "not necessary" for the state, though it does concede that for states with high numbers such a tool is "an attractive alternative." Fourthly, there are concerns about the use of historical data, especially arrest data, to make predictions that could be biased. Fifthly, the Commission points to the risk of racial bias and highlights ProPublica's case study of the use of the COMPAS algorithm to predict recidivism rates. Sixth, the report finds the proprietary nature of most tools as problematic for defendants to challenge results and for the jurisdiction as a whole. Ultimately the Commission found that implementation would not lead to a "drastic improvement" in bail decisions. (pages 9 and 10 of report).

Research resources

Government documents: Special commission on bail reform report (see pages 8-10 for risk assessment judgement).

Open letter opposing use of tool.

Media: Wired

Santa Cruz stops use of predictive policing and bans future use

Summary:

In June 2020 Santa Cruz banned the use of predictive policing. It had been one of the first cities to experiment with predictive policing, piloting and then adopting it in 2011. The Santa Cruz Police say they stopped using predictive policing in 2017. City Council voted in 2020 to ban it after years of civil society mobilization and concerns about how it perpetuated discrimination and jeopardized the safety and civil rights of residents.

Why was it dropped?

It has been reported that a coalition comprising dozens of civil liberty and racial justice groups worked together to have the technology banned. They raised concerns about how the technology contributed to discriminatory policing and jeopardizes the liberty and safety of residents. The decision to ban the technology was supported by the Santa Cruz Police Chief Andy Mills who is quoted in the Santa Cruz Sentinel as saying: "Predictive Policing has been shown over time to put officers in conflict with communities rather than working with communities."

Sources

Media: Santa Cruz Sentinel, Reuters
Civil Society: Electronic Frontier Foundation

New Jersey bans police from using Clearview AI facial recognition

Summary

In January 2020 New Jersey's Attorney General Gurbir Grewal placed a moratorium on the use of Clearview AI's facial recognition app across the state's 21 counties. This followed a *New York Times* article that made public that Clearview was scraping data from around the web, including social media sites, to amass a database of over 3 billion "publicly available" images without users' consent. The ban is temporary until guidance is drafted.

Why was it dropped?

The attorney general was alarmed by the *New York Times* article: "Until this week, I had not heard of Clearview AI," Mr. Grewal said in an interview. "I was troubled. The reporting raised questions about data privacy, about cybersecurity, about law enforcement security, about the integrity of our investigations." He said in an email interview with Law360: "[W]e need to have a sound understanding of the practices of any company whose technology we use, as well as any privacy issues associated with their technology." Grewal has also said an outright ban would be "an overcorrection that could potentially undermine public safety."

In addition to placing a moratorium on the Clearview app, the New Jersey attorney general's office has asked the state's Division of Criminal Justice to look into how state law enforcement agencies have used the app. Mr. Grewal wants to know which ones are using "this product or products like it," and what information those companies are tracking about police investigations and searches.

Further notes of interest:

Clearview AI's database appears to have broader data than that of many of its competitors. While most facial recognition programs allow law enforcement to compare images of suspects to databases composed of mug shots, driver's license photographs, and other government issued or –owned photos (and usually confined to the state in which they operate), Clearview's data appears to be national in scope and contain information from social media sites as well—like Facebook, Twitter, Venmo, YouTube and elsewhere on the Internet. All told, the database contains more than three billion photos. And it is used by more than 600 law enforcement agencies, ranging from local police departments to the F.B.I. and the Department of Homeland Security.

An Illinois a class action lawsuit was filed against Clearview AI in the same week that New Jersey placed a moratorium on the technology. Illinois citizens claim that the facial recognition app violates the Illinois Biometric Information Privacy Act (BIPA), a law that safeguards state residents from having their biometrics data used without consent.

Sources

Government documents: Cease and desist letter, open letter from Sen. Markey

Media: New York Times+

Pittsburgh predictive policing program stopped

Summary

The city of Pittsburgh announced in June 2020 that it had stopped its predictive policing program following concerns of racial bias. According to the Pittsburgh Gazette the precise date on which city officials halted the program and the extent to which they used it previously are unclear, but it was paused by Mayor Bill Peduto in December 2019. Developed by the Metro21: Smart Cities Institute at Carnegie Mellon University, the program was first piloted in 2017 and it predicted "hotspots" for criminal activity which patrol officers were then dispatched to.

Why was it dropped?

Media articles report that concerns of racial bias contributed to the decision to halt the program, especially among the members of the Pittsburgh Task Force on Public Algorithms. The Task Force had also criticised the lack of transparency and public engagement around the deployment of the program. Further, in a letter from June 2020 to the Task Force about scrapping the program, Mayor Peduto suggested that predictive analytics could be repurposed for use in social services: "Hot Spots may benefit from the aid of a social worker, service provider or outreach team, not traditional policing."

Sources

Media - Route Fifty, Pittsburgh Post-Gazette

Canada

Quayside Development Plans Stopped

Summary

In 2017 Google affiliate Sidewalk Labs and Toronto Waterfront partnered to develop plans for a smart city called Quayside. The development was highly controversial because of the ubiquity of digital technologies proposed, increasing information about how data would be collected, ownership of the data as well as concerns about oversight, citizen rights, and transparency. Concern was also raised about the large area that would be under Sidewalk Labs' control.

Why was it dropped?

The company said that Covid 19 introduced economic uncertainty making the project not financially viable. A Councillor and member of the Waterfront Toronto Board responded to the news by noting that concerns about data collection and digital governance led to reviews and significant public consultation. There was a great deal of community mobilization in opposition to the proposal that played a significant role in generating debate, interventions and research. For example, see #BlockSidewalk.

Sources:

Media: CBC, The Star, Medium
Civil Society: Block Side Walk, The Conversation

Canadian government Phoenix automated payroll system failure

Summary

A central government automated payroll system left approximately 150,000 civil service employees with incorrect or late payments (50% error rate). The Transformation of Pay Administration Initiative to update the old system started in 2009 and was meant to centralise pay operations for government workers and save taxpayers $70 million a year. Instead, a Senate Committee reported in 2018 that these savings were not realized and the system ended up costing Canadians $2.2 billion.

Why was it dropped?

In 2018 a Senate committee report (see sources) found problems with the system rollout caused "significant anxiety, stress and hardship" for thousands of employees. "By any measure, the Phoenix pay system has been a failure," the committee concluded. The Senate committee also said it was dismayed that the project went ahead with minimal independent oversight, and that no-one had accepted responsibility for the problems with the system, nor been held to account. In June 2019 a Phoenix Damages Agreement was co-developed with federal public service unions. In an interview, Alex Benay, Canada's chief information officer, said "So many things went wrong: design, procurement, project management approach, possibly culture."

Sources

Government documents: Report from Auditor General 2017 and 2018, Standing Senate Committee Report

Toronto police halt use of Clearview AI facial recognition technology
(Case study discussion in following section)

Summary
After the *New York Times* reported on the use of Clearview AI's facial recognition technology among US police forces, it emerged in Canadian media that some members of the Toronto Police Service (TPS) had been using Clearview AI in October, 2019, "with the intent of informally testing this new and evolving technology." In January 2020, Toronto police told CBC News they used facial recognition but denied using Clearview AI. TPS Chief Mark Saunders gave the order to cease using the product on February 5, 2020 after being made aware of its use.

Why was it dropped?
News that the Toronto Police were using Clearview led to controversy and concerns by civil liberty and rights groups as well as the privacy commissioner. After news of use of the system, TPS spokesperson Meaghan Gray said the force has requested that Ontario's Information and Privacy Commissioner (IPC) and the Crown Attorney's Office work with the force to review the technology's appropriateness as a tool for law enforcement, "given that it is also used by other law enforcement agencies in North America." TPS have said until a "fulsome review of the product is completed" it will not be used by the force.

Sources
Media: CBC, *Toronto Star*, *Globe and Mail*, IT World Canada

Ottawa Police Service – NEC NeoFace Reveal

Summary
Citizen Lab reports that the Ottawa Police Service ran a three-month pilot of NEC's NeoFace Reveal facial recognition system. At the end of the pilot the system was taken no further. It was reported that the the Police Service found the system to have some benefits but chose not to go forward with it. Details are not provided about why this decision was made.

Why was it rejected?
After a review following a three-month trial version it was not taken any further. It is not clear why this was not taken further.

Further notes of interest
According to Ottawa Citizen, following a series of deadly shootings in ByWard Market, Ottawa, in 2019, CCTV cameras became a hot-button issue.

Sources:
Research: Citizen Lab report
Media: Ottawa Citizen
Corporate: NEC Marketing

Vancouver Police Department – Facial recognition and driver's licence photographs

Summary

In 2012 the Insurance Corporation of British Columbia (ICBC) offered to let Vancouver police use its software to match photographs of rioters with driver's licence photographs in its database. This came to the attention of British Columbia's Privacy Commissioner, who launched an investigation into the technology and eventually ordered that this practice stop.

The Commissioner ruled that while ICBC can use the technology to detect and prevent driver's licence fraud, the corporation failed to notify its customers that facial recognition is in use. She added that the police do have the power to request personal information from ICBC but they must do it through a warrant or a court order, rather than freely between Vancouver police and ICBC.

Vancouver police did not respond to ICBC's offer in this instance and it seems that the Commissioner's investigation may have foreclosed the possibility.

Why was it rejected?

The Privacy Commissioner for British Columbia said at the time: "A public body can only use personal information for the original purpose it was collected, except in very limited circumstances. ICBC's offer to use its database to check police-submitted images is clearly a different purpose." [2]

Sources

Research: Citizen Lab report (p.62)
Media: CBC

Cleaview stops offering its facial recognition technology in Canada.

Summary

Reporting revealed a number of police forces in Canada have used Clearview AI's facial recognition system. The majority, if not all, of these examples may be related to police forces informally testing a free trial version of the software.

BuzzFeed News quote Hoan Ton-That, CEO of Clearview, in February 2020 claiming that the company was focused on doing business in the USA and Canada. As of July 6th 2020, in the wake of multiple investigations by privacy protection authorities in Canada, Clearview ceased to offer its facial recognition technology in Canada.

Police forces identified as at one time trying Clearview AI's system include: Royal Canadian Mounted Police, Calgary Police Service, Edmonton Police Service, Toronto Police Service, Peel Regional Police Service, Halton Police Service, Ottawa Police Service, Durham Regional Police Service, Niagara Regional Police Service, Hamilton Police Service, Via Rail Police Service, police in Halifax, Cornwall, and London. *The Toronto Star*, using data obtained by BuzzFeed News, identified at least 34 police forces across Canada who had obtained log-ins and searched Clearview AI's database. Investigators with the Insurance Bureau of Canada were also identified to have used the system.

Why was it rejected?

It is difficult to say but it appears that, in the face of multiple investigations by privacy authorities in Canada into Clearview AI's use, police forces moved away from the company. The Office of the Privacy Commissioner of Canada writes of a cross-agency investigation into the use of Clearview AI: "The investigation was initiated in the wake of numerous media reports that have raised questions and concerns about whether the company is collecting and using personal information without consent."

We should bear in mind, however, that facial recognition technologies are still used by police in Canada. Toronto Police Service, for example, have purchased a facial recognition system provided by NEC.

Sources

Research: Citizen Lab report
Media: Toronto Star, BuzzFeed News
Privacy Commissioners: Office of the Privacy Commissioner of Canada, Announcement of Commissioners' joint investigation

Royal Canadian Mounted Police suspends contract with Clearview AI

Summary

Of the many police forces in Canada who temporarily tried out Clearview AI's facial recognition offering, it appears that the Royal Canadian Mounted Police (RCMP) were the most advanced. We do not know how the RCMP used or planned to use the system, but some reporting by CBC quotes the RCMP as saying that "a few units in the RCMP" were using the system to "enhance criminal investigations". They also report that the force had said it had used the technology for about four months in its child exploitation unit.

The RCMP had entered into a contract with Clearview AI, which they suspended sometime around July 2020, marking the company's withdrawal from Canada.

On the 28th February 2020, the Office of the Privacy Commissioner of Canada (OPC) announced an investigation into the RCMP's use of Clearview AI's facial recognition technology. This came a week after a joint-commissioner investigation into Clearview's use across Canada was announced.

Why was it rejected?

News of the suspended contract came a number of months after the OPC's investigation was launched. There is limited information available.

Sources

Media: The Tyee, CBC
Research: Citizen Lab report
Privacy Commissioner: OPC Investigation into RCMP use of Clearview AI

Plans for use of ShotSpotter by Toronto Police Abandoned

Summary

The system was initially championed by Toronto's mayor, John Tory, and approved in the wake of a series of shootings in Toronto. When approved it had already been in use in cities across the USA. The system uses microphones to detect and locate gunfire, and automatically informs police. Questions have been raised about the effectiveness of ShotSpotter, particularly in the United States.

Plans to use the system in Toronto were abandoned "due to legal concerns". Toronto City Council "endorsed looking into the technology in July [2019], along with other measures meant to counter last summer's sharp increase in gun violence." Councillors had raised privacy concerns about the surveillance related to the possibility that the microphones could be used to eavesdrop on pedestrians.

Why was it rejected?

It seems the decision was made between councillors and the mayor; after a period of consideration it was decided not to go forward with it.

Councillor Joe Cressy is quoted in the Globe and Mail saying: "They are not proceeding for the same reason many of us voted against it in the first place ... an invasion of privacy, that there were severe risks around data collection and use ... Frankly, it was a shiny object in a RoboCop-style of enforcement model that was intended in the midst of the summer of the gun to make us all feel better."

Sources

Media: Globe and Mail, Global News article, Reveal News, WNYC

Europe

Belgium – Brussels airport drops facial recognition system

Summary:
In 2017, Zaventem Airport in Brussels silently started testing a facial recognition system. The intended full rollout of the system became public in 2019 after a newspaper published an interview with the commissioner-general of the Belgian federal police in which he mentioned the system. Following this, the Belgian Supervisory Body (autonomous federal parliamentary body in charge of monitoring the management of police information and the data protection authority for the integrated police services) made enquiries and visited the federal police service at Brussels Airport. They found out the system collected biometric data of individuals which were then compared to a self-composed "black list". This does not comply with Belgian law, so the system was stopped in 2019. More research is needed to determine if it was started again.

Why was it dropped?
An enquiry by the Supervisory Body for Police Information found the system clearly did not comply with data protection and police information law in Belgium and therefore requested the system to be (temporarily) discontinued.

Sources:
Media: Brussels Express, Teller Report, Brussels Times

Denmark Shuts Down EFI System

Summary:
EFI is a digital collection system for taxes (IT provider: Accenture) that had been planned and developed since 2005, was delayed several times and was finally introduced nationwide in 2013. However, the system was shut down in 2015 after an official review found that: "the system is generally so flawed and complex that it will be extremely difficult and very time consuming and resource consuming to fix the system errors so it can become fully functional."

Why was it dropped?
According to AlgorithmWatch's report: "The new system had serious technical as well as legal flaws and led to the loss of billions of crowns for the public, due to expired or uncollected claims". Moreover, there were several conflicts with the law, badly documented processes within the system and the knowledge lay with the IT provider (Accenture) rather than the Danish authority.

Sources:
Research: Algorithm Watch, University of Copenhagen
Media: The Local
Government: Public Accounts Committee

Denmark Scraps 'Gladsaxe model'
(Case study discussion in following section)

Summary:
The 'Gladsaxe model' was a classification system developed by the Gladsaxe municipality to trace "children who were vulnerable due to social circumstances even before they showed actual symptoms of special needs" (Algorithm Watch). The three LA's asked for exemption from the usual data protection rules in January 2018 to use the model. The government planned to make it legal for all 98 municipalities to do this. After a strong public reaction and criticism, the Liberal Alliance Government stated the proposal had been shelved in December 2018.

Why was it dropped?
The system was criticized publicly, politically and by academic researchers. Following news of a municipal data leak, the Liberal Alliance's spokeswoman Christina Egelund stated that municipalities were not equipped to deal with "the great responsibility that lies in taking care of the personal data of the citizens".

Sources:
Research: Algorithm Watch, College of Europe
Media: Politikken

France – Pilot uses of Facial Recognition in High Schools in Nice and Marseille Blocked

Summary:
Labelled an experiment, a high school in Nice and one in Marseille were equipped with facial recognition technology to grant access to school for students and to also "follow the trajectory of people." Digital and Human Rights organisations, parents and teachers' unions as well as France's National Data Protection Commission challenged these systems, with a number of groups coming together to mount a legal challenge.

Why were they dropped?
In Feb. 2020 the Administrative Court of Marseille ruled that the systems were in violation of the General Data Protection Regulation. The Court ruled that the systems breached GDPR because students are not able to give their consent freely in these circumstances and that the use of the systems was not in line with the GDPR's rules on proportionality and data minimization.

Sources:
Media: B.I.R.D, Biometricupdate.com, IAPP.org

Germany – Baden-Württemberg drops PRECOBS
(Case study discussion in the following section)

Summary:
The German federal state Baden-Württemberg stopped using the PRECOBS predictive policing system after trialling it for 4 years in the cities Stuttgart and Karlsruhe (2015-2019). This was decided after an evaluative study by the research institute *Max-Planck Institute for the Study of Crime, Security and Law* that found disappointing results in terms of reducing the number of burglaries through PRECOBS.

Why was it dropped?
The police forces in Stuttgart and Karlsruhe reported that the system was found as not efficient enough after trialling it for four years. They argue that PRECOBS is useful for statistics, but not very good at making predictions because it only captures specific types of professional burglary, while many others are not considered. Moreover, an evaluative study by the Max-Planck Institute concluded that PRECOBS's usefulness varies depending on the area it is applied in, and that for many areas, it is not very useful for predictions.

Sources:
Media: NTV, STN,GQ
Research: Max-Plank Institute for the Study of Crime, Security and Law, Evaluative Study, Research Paper, Gerstner (2018), Krasmann et al. 2019

Netherlands – Judge orders SyRI unlawful
(Case study discussion in the following section)

Summary:
The Systeem Risico Inventarisatie (System Risk Indication or SyRI) is a big data analysis system that ran under the auspices of the Ministry of Social Affairs and Employment, aiming to assess risks of welfare abuse and tax fraud. The nationwide system has been in use since 2008, was passed into law in 2014, and abandoned in 2020 after it was found to be unlawful and not compliant with Article 8 paragraph 2 of the ECHR.

Why was it dropped?
After a coalition of privacy and civil rights groups brought a lawsuit against the Dutch government's use of SyRI, the District Court of the Hague ruled that the legislation governing the deployment of SyRI violates higher law and does not comply with Article 8 of the European Convention on Human Rights (ECHR). This provision requires a fair balance (a reasonable relationship) between the social interest served by the legislation and the infringement of private life that the legislation makes.

Sources:
Research: Algorithm Watch, Bij Voorbat Verdacht, Privacy International, de Rechtspraak, PILP, UN Special Rapporteur Submission

Polish Government Ends Unemployment Scoring System

Summary:

In 2014, Poland implemented a nationwide scoring and profiling system for the unemployed as part of the wider scheme Publiczne Służby Zatrudnienia (PSZ). The ADS assigns one of three categories for all people unemployed that determines the assistance they receive. The system was heavily criticised by civil society organisations, Poland's data protection authority and the Human Rights Commissioner. Following this criticism, an official review by the Supreme Audit Office took place (which, among other findings, found 80% of staff to be unhappy with the system) and the Human Rights Commissioner referred the case to Poland's Constitutional Court due to a lacking legal base of the system. The court ruled the system to be a breach of Poland's constitution in 2018 and the government decided to end the system in 2019.

Why was it dropped?

The system received a great deal of criticism from civil society, Poland's data protection authority and the Human Rights Commissioner about its lack of transparency, lack of oversight, and the alleged arbitrary nature of decision-making due to the simplification of data, arguing that the system infringes data protection. After a negative official review and a court case, Poland's Constitutional Court ruled that the system breaches Poland's constitution.

Sources:

Research: Algorithm Watch, Jedrzej Niklas (2019) (2018),

Slovenia Cancels Automatic License Plate Recognition

Summary:

Following new police legislation, ZNPPol-A, police in Slovenia started to pilot an automatic optical license plate recognition system in 2017. After a formal complaint by Slovenian Human Rights Ombudsman and the Information Commissioner, the Constitutional Court ruled in 2019 that the automatic license plate recognition contravenes the constitutional right to protection of personal data. Two other parts of the complaint were air passenger data (see below) and police use of drones (which was allowed).

Why was it dropped?

The system was heavily criticised by the Slovenian Human Rights Ombudsman and the Information Commissioner, who also filed a formal complaint. They argued that the new police legislation contradicts constitutional and convention standards on the protection of privacy in three areas (air passenger data – see below; police use of drones – was allowed). The Constitutional Court then ruled that automatic license plate recognition contravenes the constitutional right to protection of personal data.

Research resources:

Media: Slovenia News, Slovenia News
Government: Ombudsman Report,

Swedish Government Cancels Automated Welfare Payments System

Summary:
The Swedish Government started using an automated system to check that people receiving a certain type of unemployment benefit keep up their obligations and issue warnings or withhold payments if not. The nationwide system was switched off in 2018.

Why was it dropped?
A review in 2018 found that 10-15% of the decisions made by the system were wrong and the system was then dropped. Media reports indicate that case worker protest was a factor in signalling the problems with this system.

Sources:
Research: Algorithm Watch
Media: BreakIt

Switzerland – Solothurn cancels PRECOBS

Summary:
The police in Solothurn considered using the predictive policing system PRECOBS that is already in use by other police forces in Europe and also in other places in Switzerland (Aargau, Zürich and Baselland). It is a location-based system that aims to prevent burglary. Solothurn police observed the use in two other cities, evaluated the system carefully and finally concluded their evaluation with a negative assessment.

Why was it dropped?
After evaluating the system, Solothurn police decided against implementing the system. They found the system to be not useful enough, that no clear benefit could be identified, and that the system would not lead to savings and was too expensive for their financially strained budget.

Sources:
Media: Solothurner Zeitung, Schweiz, SRF
Online source: Wikipedia

United Kingdom – Government Scraps Use of Algorithm to Determine A Level Results

Summary:

Students across the UK and Wales were not able to sit their A-Level exams in 2020 due to Covid 19. These exams are crucial for university access. The Office of Qualifications and Examinations Regulations (Ofqual) introduced an algorithm to predict the results that students would 'likely' have achieved based on the historical distribution of grades in a school, teacher's predictions, predicted grade based on school grades in the past and class size. When the grades were released many students received grades lower than expected. It's been reported that in England 40% of students had their grades downgraded from their teachers' assessments.

Why was it dropped?

The government faced public outrage as students and parents protested as the algorithm was widely criticized as being classist and inaccurate. As reported, the algorithm was designed to mirror results from past years which meant also bakes in biases and also limits the ability for schools and students to perform by "locking them in." Legal action was also threatened by Foxglove, a digital rights organisation. The Office for Statistics Regulation is now conducting a review of the approach taken by Ofqual.

Sources:

Media: Guardian, BBC, the Verge, StateWatch

Research: LSE Blog

Durham Police cancels HART (and Mosaic)
(Case study discussion in following section)

Summary:
One of the first algorithmic risk assessment tools to be used by a UK police force, development of HART started in 2013 and it was implemented in mid-2016. HART was the result of a collaboration between Durham Police and statistical experts at Cambridge University and formed part of an ongoing partnership between the two institutions. HART was created to help officers decide which individuals could be referred to Durham Police's rehabilitation program, Checkpoint, designed to find alternatives to prosecution. Problems with accuracy were found in a 2016 validation study of the model, in that the tool was better at predicting low risk offenders than high risk offenders, meaning that in order to compensate for the possibility of wrongly predicting a high risk offender to be low risk the HART algorithm overestimated high risk predictions.

HART also controversially made use of Experian's Mosaic segmentation tool which was revealed in a 2018 investigation by Big Brother Watch. However one of the HART project leaders, Sheena Urwin (Head of Criminal Justice at Durham Police), said the force stopped using Mosaic data in 2018, citing financial concerns rather than ethical concerns as a reason for this.

Why was it dropped?
Durham's Freedom of Information department confirmed that HART was dropped in September 2020, but did not disclose who gave the instruction to cancel use or why except that "the model was no longer being supported. The HART triage tool was only ever used to assist the custody officer in his decision-making process on whether an individual detained would be eligible for the Checkpoint scheme. At no point was it used as the sole decision-making mechanism."

Within Durham Constabulary there was considerable awareness of HART's potential for bias and steps were taken to try mitigate this: for example an ALGOCARE ethical framework to guide HART decisions was introduced and custody officers attended awareness sessions relating to unconscious bias. At the same time, some civil society groups like Big Brother Watch have been critical about the inclusion of postcode data, as well as the Constabulary's use of Experian's Mosaic segmentation tool. It's also worth noting that then-Chief of Police Mike Barton told the *Financial Times* that HART would be cancelled if it did not work but that he thought it was worth testing it.

Sources:
Media: Wired, Financial Times
Civil society: Big Brother Watch
Academia: Cambridge University, Oswald et al review of HART

Hackney Council drops Xantura' EHPS
(Case study discussion in the following section)

Summary:
In 2015, Hackney Council started using Xantura's Early Help Profiling System (EHPS) to risk assess at-risk families. The predictive system would send an alert and a report to case workers if a risk threshold had been crossed. The pilot scheme was dropped in 2019.

Why was it dropped?
A Hackney Council spokesperson is quoted in the Hackney Citizen as saying: "At the conclusion of the pilot we had not been able to realise the expected benefits and decided to not continue beyond the pilot stage. We found that the data available was more limited than had initially been envisaged and issues of variable data quality meant that the system wasn't able to provide sufficiently useful insights to justify further investment in the project." Concerns were raised about privacy and consent in media coverage as well as by a local politician.

Sources:
Media: Guardian, Hackney Citizen, Community Care,

Kent Police Cancels PredPol

Summary:
Kent Police were the first force in the UK to introduce predictive policing in 2013 when they started using PredPol – a location-based system to prevent crime. In 2018, they ended the project. They are now considering developing their own spatial mapping system.

Why was it dropped?
The force reported that they used PredPol as a preventative tool rather than for prediction as it "did not predict crime". They also said that the system "had a good record of predicting where crimes are likely to take place", it was "more challenging to show that we have been able to reduce crime with that information". Moreover, Kent police force wants to "avoid paying an ongoing licence fee to an external company".

Sources:
Media: New Scientist, BBC, FT, Telegraph

North Tyneside Council stops using RBV
(Case study discussion of RBV in following section)

Summary:
Like many other UK councils, North Tyneside Council started using the RBV system after a "Housing Benefit and Council Tax Benefit Circular" by the Department for Work and Pensions. The 2011 document "outlines the Department's policy on Risk-Based Verification (RBV) of Housing Benefit and Council Tax Benefit (HB/CTB) claims" and expressed the Department's wish "to extend RBV on a voluntary basis to all LAs [local authorities] from April 2012". RBV is described as a system that applies different levels of checks to benefit claims and "assigns a risk rating to each HB/CTB claim". North Tyneside Council started using RBV through the CallCredit, now TransUnion, system at some point after 2015 and stopped using it in 2019.

Why was it dropped?
They dropped the system because it often wrongly identified low-risk claims as high-risk and did not give a reason for the high-risk categorisation, thereby delaying payment of welfare to an unknown number of people.

Sources:
Media: Guardian,
Government: North Tyneside Council Report, North Tyneside Report to Cabinet, North Tyneside RBV Policy, Housing and Benefit Tax Circular

Two Councils Drop Use of RBV

Two other UK local councils have dropped their uses of RBV, this information comes from anonymous interviews as part of the Data Justice Lab 'Towards Democratic Auditing' project.

Summary:
Like many other councils, these councils started using the RBV system after a "Housing Benefit and Council Tax Benefit Circular" by the Department for Work and Pensions. The 2011 document "outlines the Department's policy on Risk-Based Verification (RBV) of Housing Benefit and Council Tax Benefit (HB/CTB) claims" and expressed the Department's wish "to extend RBV on a voluntary basis to all LAs [local authorities] from April 2012". RBV is described as a system that applies different levels of checks to benefit claims and "assigns a risk rating to each HB/CTB claim". These councils started using the RBV system in 2017 and in 2019 but dropped the system in 2019 and 2020

Why was it dropped?
One council representative explained that the system did not make a material difference (neither in time savings nor in, for example, postal savings), that it was not reducing contact and that the council was "well within our processing times" anyway. Moreover, they explained that the rollout of the Universal Credit System meant the actual housing benefit caseload would be decreasing soon anyway. Another interviewee explained that their council wanted to drop the system because the software is expensive, they felt they were "not getting the best use out of it" and needed to save money.

Bristol Stops Using RBV

Summary:

Bristol provides an overview of its use of RBV and also details why it decided to not renew its contract with its RBV supplier as of September 2020.
It's reported that the Council has been using the approved tool since October 2014. The tool is used to process new claims. The system is an external system that classifies the level of risk as low, medium and high. Low risk cases require little verification while more checking is required when someone is ranked high risk. The categorizations of risk "are confidential to the supplier so local authorities do not know why cases are categorised as they are."

Why was it dropped?

It's noted that "the system has not delivered the anticipated savings in workload for staff or significant improvements in average processing times."

Sunderland City Council discontinues Palantir Intelligence Hub

Summary:

In 2014, Sunderland City Council started using Palantir's "intelligence hub" to bring together data of the Troubled Families Programme and also to help find areas at risk of flooding. After using the system for 5 years, the contract with Palantir ended in 2019 and the Council has not renewed the contract. It is unclear if this was planned all along or was a deliberate decision against Palantir's system.

Why was it dropped?

Sunderland City Council has not renewed Palantir's contract and wants to become "self-sufficient" and build an in-house system instead. They argue that all contractual obligations have been completed and a full skills transfer has taken place.

Sources:

Media: Guardian,
Civil Society: Sunderland for Transparency, Nesta

Home Office - visa

Summary:

The Home Office started using an automated tool to sort visa applications by ranking them green, amber and red. The Joint Council for the Welfare of Immigrants and Foxglove, a technology rights group, argued that the algorithm discriminated on the basis of nationality and took the Home Office to court.

Why was it dropped?

The Home Office announced that it would stop using the algorithmic ranking system in August 2020 before the judicial review. The Home Office said they would be reviewing and redesigning their processes.

Sources:

Media: Guardian, BBC

Civil Society: JCWI, Foxglove

West Midlands police discontinue Most Serious Violence Tool

Summary

The Most Serious Violence (MSV) Tool was part of the government funded National Data Analytics Solution (NDAS) project, which has received at least £10 million during 2018-2020 from the Home Office. Designed to predict gun and knife crime before it happens, the MSV tool was trained on data from crime and custody records of 3.5 million people living in the West Midlands and West Yorkshire, intelligence reports, and the Police National Computer database. The MSV tool was tested but never used in active policing.

Why was it discontinued?

Documents published by the West Midlands' Police Ethics Committee, reported by Wired, state the tool was found to be inaccurate: "it has proven unfeasible with data currently available, to identify a point of intervention before a person commits their first MSV offense with a gun or knife, with any degree of precision." NDAS had claimed the tool was up to 75 percent accurate but in testing the tool the West Midlands Police found the accuracy to be 14 to 19 percent for the West Midlands, and 9 to 18 percent for West Yorkshire. According to Wired a coding "flaw" within the tool made it incapable of accurately predicting violence. Tweaks and fixes brought the best-case accuracy up to 25-38 percent for the West Midlands and 36-51 percent for West Yorkshire Police, but the proposal to continue developing the tool was ultimately rejected by the West Midlands Police ethics committee in August 2020. (The committee is a voluntary group consisting of experts from different fields.)

There have also been concerns raised about the type of data the system used as "predictors" having the potential to produce biased results.

Sources

Media - Wired
Corporate - AI Botics

Australia

LINK Crime Database

Summary
The LINK crime database system was intended to link police information to help produce insights. It seems it was dropped because of poor management in the attempted deployment, particularly with difficulties in linking it to existing systems. LINK was intended to replace an earlier system, LEAP.

Why was it dropped?
It was reported that the system was suspended in 2011 because of the difficulties and the cost of linking the off the shelf system with other systems operated by the Victoria police.

Sources:
Research: IEEE
Media: Computer World, IT News

Federal Police Investigations System

Summary
The Australian Federal Police had a contract with Elbit Systems to develop an "intelligence focused" case management tool. Contracts were agreed in 2013 and abandoned in 2015.

Why was it dropped?
It is not clear, with some mention of the project taking longer than expected, there are other other comments linked to functionality and cost.

Sources:
Media: IT News (2015), IT News (2018)

HealthSMART - Victoria

Summary

The goal was to develop a Victoria-wide electronic health record system that combined health-related financial systems with patient record management systems. which failed because of unforeseen costs resulting from adapting an off-the-shelf system for a local context. The project was to have brought hospitals a new clinical, patient and client management, resource management and picture archiving system. The Victoria government attempted to deploy the system between 2003-2012, with the original completion date set for June 2007. The project was abandoned in 2012.

Why was it dropped?

It was reported that the gap between the local requirements and the American system which was purchased were greater than initially thought. In the end, the base system had to be largely rewritten which led to the huge cost overruns. The system was supposed to cost $240 m AUD but ballooned to $566 m.

Sources:

Media: IT News
Research: IEEE
Government: Victoria Government, HealthSmart Victoria

Robo-debt / Online Compliance Intervention Paused
(Case study discussion in the following section)

Summary
The Department of Human Services and Centrelink, which manages social security payments in Australia, changed the algorithmic system it used to identify people who had been overpaid benefits in 2016. A new accounting method meant that fortnightly earnings could be used to estimate annual earnings. Also, where previously a human would be responsible for checking a case that was flagged through data matching, the new system was automated which meant debt notices were sent automatically to those who were identified by the system as being overpaid benefits. Thousands more letters were being sent out than had previously been the case. The system went from sending out 20,000 debt notices a year to, at one point, 20,000 a week. Further, it became the responsibility of those receiving debt notices to challenge the notification and prove that the government's system had made a mistake.

Individuals have 21 days to correct discrepancies logged by the system. Many people were missing the letters because of things like letters being sent to old addresses or posted in myGov accounts that people don't check anymore. This resulted in many people learning they had been flagged by the system when they were contacted by a debt collector [1].

Also, challenging a debt notice was difficult as it involved trying to get an officer on the phone which could take hours and finding paperwork, sometimes going back 7 years. Those receiving the letters reported their frustration on a website set up by a coalition to challenge the new system called NotMyDebt. One social service organisation reported that a quarter of the debt notices it investigated were wrong. The automated debt recovery system was labelled RoboDebt by critics.

Why was it paused?
The system faced considerable pressure from civil society organisations, activists and politicians. Victoria Legal Aid challenged the system in court and in November 2019 a judge ruled that the system was unlawful. The government paused the automated debt matching component in November. In May 2020 the Government announced it would repay all debts at a cost of $721 million. A Senate Committee has recommended the system be terminated. A class action lawsuit was settled for $1.2 billion in November 2020. It is unclear if the system will be started again in a different form.

Sources:
Media: ABC, Guardian Guardian Guardian (2019), Guardian (2020)
Research: Carney 2018
Civil Society: #NotMyDebt
Government: Ombudsman Review

New Zealand

Immigration tool to profile over-stayers

Summary
Immigration New Zealand began piloting a modelling tool to assess who might overstay their time in NZ, lead to high health costs and commit crime. Media reporting in 2018 identified that the program had been running for 18 months. The programme was criticized as targeting Indian students.

Why was it stopped?
The programme was put on hold in 2018. Media coverage suggested the system was paused because of concerns about racial profiling, Immigration NZ denied that the system considered ethnicity. The planned suspension was meant to give Immigration NZ a chance to discuss the matter with the Human Rights Commission and Privacy Commissioner John Edwards. This was in April 2018. These two later said they would work with Immigration NZ if it wanted to develop such tools, but it was reported in July 2018 that Immigration NZ was scrapping this work.

Sources:
Media: RNZ, RNZ (2), News Hub, Pundit

Vulnerable Children PRM (Predictive Risk Modelling) (NZ)
(Case study discussion in following section)

Summary
The New Zealand government began exploring the potential of using a predictive risk modelling tool to identify children with the highest risk of neglect and abuse in 2012. As noted by Philip Gillingham, the system was developed to combine multiple datasets and to identify the children in families with parents claiming public welfare benefits who were most at risk of abuse and neglect. The stated aim was to provide supportive services to families.

Why was it dropped?
The system was trialled but never implemented. There was a great deal of critique as researchers and academics reviewed the system and raised concerns about how it could further embed bias and lead to greater inequality, in addition to concerns raised about accuracy and privacy. It was cancelled by Minister of Social Development Anne Tolley before an observational study was due to take place.

Sources:
Media: RNZ
Research: Keddell (2014), Gillingham (2019), Gillingham (2016)
Government: Ministry of Social Development, Privacy Commissioner

New Zealand Police Decide Not to Pursue Use of Clearview AI Facial Recognition Technology
(Case study discussion in following section)

Summary
New Zealand's High Tech Crime Group decided to test a trial copy of Clearview's Facial Recognition system. Some testing took place with some images uploaded to the system between January and March 2020. The public and it seems other government officials learned of the trial following reporting by RNZ.

Why was it not pursued?
It was reported that the test led to a disappointing match rate and so the High Tech Crime Group decided not to pursue the technology. After media coverage of the trial the Police Commissioner ordered a stocktake of uses of surveillance technologies.

Note:
In August 2020 it was reported that police in New Zealand were setting up a facial recognition system to identify people from CCTV feeds.

Police use of facial recognition technologies is in the news again in September and October of 2020 as it has been reported that the Department of Internal Affairs has signed a master agreement with DXC Technology which enables public and private pursuit of uses of facial recognition.

Sources:
Media: RNZ, RNZ, TechWire Asia, NZ Herald

Bibliography

Adams, R., Weale, S. and Barr, C. (2020) *A-Level Results: Almost 40% of Teacher Assessments in England Downgraded, the Guardian*. Available at: http://www.theguardian.com/education/2020/aug/13/almost-40-of-english-students-have-a-level-results-downgraded (Accessed: 15 July 2021).

AI Now Institute Report (2018) *Litigating Algorithms: Challenging Government Use of Algorithmic Decision Systems*, p. 17. Available at: https://ainowinstitute.org/litigatingalgorithms.pdf.

Alexander, R. (2017) *Spokane Judges Have a New Tool to Decide Whether the Right People are in Jail, The Spokesman-Review*. Available at: https://www.spokesman.com/stories/2017/jan/11/spokane-judges-have-a-new-tool-to-decide-whether-t/ (Accessed: 7 July 2021).

Algorithm Watch (2019) 'Automating Society', *AlgorithmWatch*. Available at: https://algorithmwatch.org/en/automating-society/ (Accessed: 18 August 2021).

Allen, K. (2020) 'Toronto Police Chief Halts Use of Controversial Facial Recognition Tool', thestar.com, 13 February. Available at: https://www.thestar.com/news/gta/2020/02/13/toronto-police-used-clearview-ai-an-incredibly-controversial-facial-recognition-tool.html (Accessed: 8 July 2021).

Allen, K. and Gillis, W. (2019) 'Toronto Police have been Using Facial Recognition Technology for more than a Year', thestar.com, 28 May. Available at: https://www.thestar.com/news/gta/2019/05/28/toronto-police-chief-releases-report-on-use-of-facial-recognition-technology.html (Accessed: 8 July 2021).

Allen, K., Gillis, W. and Boutilier, A. (2020) *Facial Recognition App Clearview AI Has Been Used Far More Widely In Canada Than Previously Known*, thestar.com. Available at: https://www.thestar.com/news/canada/2020/02/27/facial-recognition-app-clearview-ai-has-been-used-far-more-widely-in-canada-than-previously-known.html (Accessed: 8 July 2021).

Alston, P. (2018) *Statement on Visit to the United Kingdom, by Professor Philip Alston, United Nations Special Rapporteur on extreme poverty and human rights, United Nations Human Rights Office of the High Commissioner*. Available at: https://www.ohchr.org/en/NewsEvents/Pages/DisplayNews.aspx?NewsID=23881 (Accessed: 18 August 2021).

Altermatt, S. (2017) *Kein klarer Nutzen: Solothurner Polizei verzichtet auf Einbruchssoftware, Solothurner Zeitung*. Available at: https://www.solothurnerzeitung.ch/solothurn/kanton-solothurn/kein-klarer-nutzen-solothurner-polizei-verzichtet-auf-einbruchssoftware-ld.1471709 (Accessed: 15 July 2021).

American Federation of Teachers (2017) *Federal Suit Settlement: End of Value-Added Measures in Houston*. Available at: https://www.educationviews.org/federal-suit-settlement-value-added-measures-houston/, https://www.educationviews.org/federal-suit-settlement-value-added-measures-houston/ (Accessed: 29 June 2021).

Amrein-Beardsley, A. (2019) 'The Education Value-Added Assessment System (EVAAS) on Trial: A Precedent-Setting Lawsuit with Implications for Policy and Practice', *eJournal of Education Policy*.

van Ark, T. (2020) 'Letter of 23 April 2020 to the President of the House of Representatives by the State Secretary for Social Affairs and Employment, Tamara van Ark, on a Court Judgment Regarding SyRI'.

Bailey, L. (2017) 'Software Firms Sued Over Mich. Jobless System', *Courthouse News Service*, 3 March. Available at: https://www.courthousenews.com/software-firms-sued-michigan-jobless-system/ (Accessed: 29 June 2021).

Barbaschow, A. (2019) *Government Backflip As Robo-Debt Income Automation Paused*, *ZDNet*. Available at: https://www.zdnet.com/article/government-backflip-as-robo-debt-income-automation-paused/ (Accessed: 12 August 2021).

Barrett, M. (2017) *After Years of Scandals, Here's How Denmark's Reformed Tax Authority Will Look*, *The Local*. Available at: https://www.thelocal.dk/20170614/heres-how-denmarks-reformed-tax-administration-will-look/ (Accessed: 14 July 2021).

BBC News (2018) 'Kent Police Stop Using Crime Predicting Software', *BBC News*, 26 November. Available at: https://www.bbc.com/news/uk-england-kent-46345717 (Accessed: 15 July 2021).

BBC News (2020) 'Home Office Drops "Racist" Algorithm from Visa Decisions', *BBC News*, 4 August. Available at: https://www.bbc.com/news/technology-53650758 (Accessed: 11 August 2021).

Bertsimas, D. *et al.* (2020) 'Bus Routing Optimization Helps Boston Public Schools Design Better Policies', *INFORMS Journal on Applied Analytics*, 50(1), pp. 37–49. doi: 10.1287/inte.2019.1015.

Big Brother Watch (2018) *A Closer Look at Experian Big Data and Artificial Intelligence in Durham Police*, *Big Brother Watch*. Available at: https://bigbrotherwatch.org.uk/2018/04/a-closer-look-at-experian-big-data-and-artificial-intelligence-in-durham-police/#_ftn2 (Accessed: 15 July 2021).

Bij Voorbaat Verdacht (2019) 'Save the Date: Datum zitting bodemprocedure tegen SyRI op 29 oktober 2019', *Bij Voorbaat Verdacht*, 23 April. Available at: https://bijvoorbaatverdacht.nl/save-the-date-zitting-rechtszaak-tegen-syri-op-29-oktober-2019/ (Accessed: 14 July 2021).

Bonnett, G. (2018a) *Immigration Dumps Controversial Deportation Analytical Tool*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/361199/immigration-dumps-controversial-deportation-analytical-tool (Accessed: 12 August 2021).

Bonnett, G. (2018b) *Immigration NZ Using Data System to Predict Likely Troublemakers*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/354135/immigration-nz-using-data-system-to-predict-likely-troublemakers (Accessed: 12 August 2021).

Bracewell-Worrall, A. (2018) 'Immigration Minister Puts "Ethnic Profiling" Pilot Programme on Hold', *Newshub*, 4 September. Available at: https://www.newshub.co.nz/home/politics/2018/04/immigration-minister-puts-ethnic-profiling-pilot-programme-on-hold.html (Accessed: 12 August 2021).

Brandimarte, E. (2019) *Should We Let the Algorithms Decide? A Critical Assessment of Automated Decision-Making in Europe*. College of Europe.

Bristol Government (2020) *Review and Rationale for Replacement of RBV*.

Brusselsexpress.eu (2019) *Facial Recognition Cameras at Brussels Airport "Have No Legal Basis", Says Lawyer*, *Brussels Express*. Available at: https://brussels-express.eu/facial-recognition-cameras-at-brussels-airport-have-no-legal-basis-says-lawyer/ (Accessed: 8 July 2021).

Burgess, M. (2018) *UK Police are Using AI to Make Custodial Decisions – But It Could Be Discriminating Against the Poor*, *WIRED UK*. Available at: https://www.wired.co.uk/article/police-ai-uk-durham-hart-checkpoint-algorithm-edit (Accessed: 15 July 2021).

Burgess, M. (2020) 'Police built an AI to predict violent crime. It was seriously flawed', *Wired UK*, 8 June. Available at: https://www.wired.co.uk/article/police-violence-prediction-ndas (Accessed: 12 August 2021).

'California Governor Signs Landmark Bill Halting Facial Recognition on Police Body Cams' (2019). ACLU of Northern CA. Available at: https://www.aclunc.org/news/california-governor-signs-landmark-bill-halting-facial-recognition-police-body-cams (Accessed: 7 July 2021).

Callahan, R. (2018) *Appeals Court Affirms Ruling That IBM Owes Indiana $78M*, *AP News*. Available at: https://apnews.com/article/36ba0562a02142e5adfe39518e2e0f85 (Accessed: 4 August 2021).

Cardozo, T. and Freeze, C. (2020) 'Toronto Police Chief Orders Officers to Stop using Clearview AI Software', 13 February. Available at: https://www.theglobeandmail.com/canada/toronto/article-toronto-police-chief-orders-officers-to-stop-using-clearview-ai/ (Accessed: 8 July 2021).

Carney, B. (2020) *FOI Documents Confirm RCMP Falsely Denied Using Facial Recognition Software*, *The Tyee*. The Tyee. Available at: https://thetyee.ca/News/2020/04/16/RCMP-FOI-Facial-Recognition/ (Accessed: 8 July 2021).

Carney, T. (2019) 'Robo-Debt Illegality: The Seven Veils of Failed Guarantees of the Rule of Law?', *Alternative Law Journal*, 44(1), pp. 4–10. doi: 10.1177/1037969X18815913.

CBC News (2012) *Police Can't Use ICBC Facial Recognition to Track Rioters*, *CBC*. Available at: https://www.cbc.ca/news/canada/british-columbia/police-can-t-use-icbc-facial-recognition-to-track-rioters-1.1207398 (Accessed: 8 July 2021).

CBC News (2020) *Toronto Police Admit Using Secretive Facial Recognition Technology Clearview AI*, *CBC*. Available at: https://www.cbc.ca/news/canada/toronto/toronto-police-clearview-ai-1.5462785 (Accessed: 8 July 2021).

Chan, J. B. (2016) *ShotSpotter Not Exactly Taking a Bite Out Of Crime*, *Reveal*. Available at: http://revealnews.org/blog/shotspotter-not-exactly-taking-a-bite-out-of-crime/ (Accessed: 8 July 2021).

Charette, R. N. (2012) *Troubled HealthSMART System Finally Cancelled in Victoria Australia*, *IEEE Spectrum*. Available at: https://spectrum.ieee.org/troubled-healthsmart-system-finally-cancelled-in-victoria-australia- (Accessed: 12 August 2021).

Charette, R. N. (2018a) *Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold*,

IEEE Spectrum. Available at: https://spectrum.ieee.org/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold (Accessed: 4 August 2021).

Charette, R. N. (2018b) Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold - IEEE Spectrum, IEEE Spectrum: Technology, Engineering, and Science News. Available at: https://spectrum.ieee.org/riskfactor/computing/software/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold (Accessed: 29 June 2021).

Chiusi, F. (2020) Automating Society Report 2020. Berlin, Germany: AlgorithmWatch, p. 298. Available at: https://automatingsociety.algorithmwatch.org (Accessed: 14 July 2021).

Chowdhury, H. (2018) Kent Police Stop Using Crime Predicting Software, The Telegraph. Available at: https://www.telegraph.co.uk/technology/2018/11/27/kent-police-stop-using-crime-predicting-software/ (Accessed: 15 July 2021).

Clarke, T. (2010) *Victorian Police Suspend Crime Database Overhaul*, *Computerworld*. Available at: https://www.computerworld.com/article/3462991/victorian-police-suspend-crime-database-overhaul.html (Accessed: 12 August 2021).

'CNIL Bans High Schools' Facial-Recognition Programs' (2019) *iapp*, 29 October. Available at: https://iapp.org/news/a/cnil-bans-high-school-facial-recognition-programs/ (Accessed: 14 July 2021).

Coleman, E. (2020) *One City Rejected a Policing Algorithm. Could It Be Used For Other Purposes?*, *Route Fifty*. Available at: https://www.route-fifty.com/public-safety/2020/07/pittsburgh-hot-spot-algorithm/166953/ (Accessed: 8 July 2021).

Commonwealth Ombudsman (2019) *Submission by the Commonwealth Ombudsman: Inquiry Into Centrelink's Compliance Program*.

Conger, K., Fausset, R. and Kovaleski, S. F. (2019) 'San Francisco Bans Facial Recognition Technology', *The New York Times*, 14 May. Available at: https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html (Accessed: 7 July 2021).

Coop, A. (2020) *Privacy Expert Calls Toronto Police's Use of Clearview AI 'Appalling'*, *IT World Canada News*. Available at: https://www.itworldcanada.com/article/privacy-expert-calls-toronto-polices-use-of-clearview-ai-appalling/427346 (Accessed: 8 July 2021).

Coughlan, S. (2020) 'Why did the A-level algorithm say no?', *BBC News*, 14 August. Available at: https://www.bbc.com/news/education-53787203 (Accessed: 15 July 2021).

Counsel to the Attorney General (2020) 'Cease and Desist'.

County of Los Angeles Office of Child Protection (2016) *Examination of Using Structured Decision Making and Predictive Analytics in Assessing Safety and Risk in Child Welfare*. County of Los Angeles, Office of Child Protection, p. 18.

Cournoyer, C. (2017) *Too Much Data? Illinois Abandons System Meant to Predict Child Abuse*, *Governing: The Fu*

ture of States and Localities. Available at: https://www.governing.com/archive/tns-chicago-data-mining.html (Accessed: 29 June 2021).

Cowan, P. (2015) *AFP Dumps $145m Investigation System Replacement*, *iTnews*. Available at: https://www.itnews.com.au/news/afp-dumps-145m-investigation-system-replacement-408034 (Accessed: 12 August 2021).

Cowie, T. (2015) *Insight for 21 June 2015 - Child Abuse or Big Brother?*, *RNZ*. Available at: https://www.rnz.co.nz/national/programmes/insight/audio/201758628/insight-for-21-june-2015-child-abuse-or-big-brother (Accessed: 13 August 2021).

Crenshaw, K. (2015) 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', *University of Chicago Legal Forum*, 1989(1). Available at: https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8.

Crist, B. (2019) 'An Important Update from Chef', *Chef Blog*, 23 September. Available at: https://blog.chef.io/an-important-update-from-chef (Accessed: 7 July 2021).

Crozier, R. (2012) *Victoria kills HealthSMART IT project*, *iTnews*. Available at: https://www.itnews.com.au/news/victoria-kills-healthsmart-it-project-301410 (Accessed: 12 August 2021).

Data Justice Lab (2018) *Towards Democratic Auditing: Civic Participation in the Scoring Society*, *Data Justice Lab*. Available at: https://datajusticelab.org/towards-democratic-auditing-civic-participation-in-the-scoring-society/ (Accessed: 11 August 2021).

'David Mutnick, for himself and others similarly situated, vs. Clearview AI, Inc., Hoan Ton-That and Richard Schwartz' (2020). US District Court for the Northern District of Illinois Eastern Division. Available at: https://www.scribd.com/document/444154093/gov-uscourts-ilnd-372790-1-0 (Accessed: 8 July 2021).

de Rechtspraak (2020) *SyRI-wetgeving in strijd met het Europees Verdrag voor de Rechten voor de Mens*, *de Rechtspraak*. Available at: https://www.rechtspraak.nl/Organisatie-en-contact/Organisatie/Rechtbanken/Rechtbank-Den-Haag/Nieuws/Paginas/SyRI-wetgeving-in-strijd-met-het-Europees-Verdrag-voor-de-Rechten-voor-de-Mens.aspx (Accessed: 14 July 2021).

Department for Work and Pensions (2011) *Housing Benefit and Council Tax Benefit Circular*. HB/CTB S11/2011. London.

Devanesan, J. (2020) *New Zealand Police Trial Controversial Clearview AI Facial Recognition Tech*, *Tech Wire Asia*. Available at: https://techwireasia.com/2020/05/police-trial-controversial-clearview-ai-facial-recognition-tech-in-new-zealand/ (Accessed: 13 August 2021).

Digital Society Blog (2018) 'Profiling the Unemployed', *HIIG*, 16 January. Available at: https://www.hiig.de/en/profiling-the-unemployed/ (Accessed: 15 July 2021).

Eaton, M. and Bertoncin, C. (2018) *State of Offices of Data Analytics (ODA) in the UK*. Nesta.

Edes, A. and Bowman, E. (2018) *'Automating Inequality': Algorithms In Public Services Often Fail The Most Vulnerable*, NPR.org. Available at: https://www.npr.org/sections/alltechconsidered/2018/02/19/586387119/automating-inequality-algorithms-in-public-services-often-fail-the-most-vulnerab (Accessed: 29 June 2021).

Egbert, S. and Krasmann, S. (2019) Predictive Policing. Eine ethnographische Studie neuer Technologien zur Vorhersage von Straftaten und ihre Folgen für die polizeiliche Praxis. Hamburg: Universität Hamburg.

Eubanks, V. (2015) 'Want to Cut Welfare? There's an App for That', 27 May. Available at: https://www.thenation.com/article/archive/want-cut-welfare-theres-app/ (Accessed: 29 June 2021).

Eubanks, V. (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. First Edition. New York: St Martin's Press.

European Court of Human Rights (2020) *European Convention on Human Rights*. France: Council of Europe, p. 34.

Final Report of the Special Commission to Evaluate Policies and Procedures Related to the Current Bail System (2019). Government document. Boston: The General Court of Massachusetts State House, p. 114.

Finley, K. (2019) 'Software Company Chef Won't Renew ICE Contact After All', *Wired*, 23 September. Available at: https://www.wired.com/story/software-company-chef-wont-renew-ice-contact/ (Accessed: 7 July 2021).

Foxglove (2020) *Home Office Says it Will Abandon its Racist Visa Algorithm - After We Sued Them*, *Foxglove*. Available at: https://www.foxglove.org.uk/2020/08/04/home-office-says-it-will-abandon-its-racist-visa-algorithm-after-we-sued-them/ (Accessed: 11 August 2021).

Galindo, G. (2019) *'No legal basis' for facial recognition cameras at Brussels Airport*, *The Brussels Times*. Available at: https://www.brusselstimes.com/brussels-2/60362/no-legal-basis-for-facial-recognition-cameras-identity-brussels-airport-intelligent-cameras/ (Accessed: 8 July 2021).

Gavaghan, C. (2018) *Where Did it Algo wrong? The Threat and Promise of Predictive Analytics*, *Pundit*. Available at: https://www.pundit.co.nz/content/where-did-it-algo-wrong-the-threat-and-promise-of-predictive-analytics (Accessed: 12 August 2021).

Gerstner, D. (2017) *Predictive Policing als Instrument zur Prävention von Wohnungseinbruchdiebstahl: Evaluationsergebnisse zum baden-württembergischen Pilotprojekt P4*. Freiburg im Breisgau: Max-Planck-Institut für ausländisches und internationales Strafrecht.

Gerstner, D. (2018) 'Predictive Policing in the Context of Residential Burglary: An Empirical Illustration on the Basis of a Pilot Project in Baden-Württemberg, Germany', *European Journal for Security Research*, 3(2), pp. 115–138. doi: 10.1007/s41125-018-0033-0.

Ghaffary, S. (2019) *San Francisco's Facial Recognition Technology Ban, Explained*, *Vox*. Available at: https://www.vox.com/recode/2019/5/14/18623897/san-francisco-facial-recognition-ban-explained (Accessed: 7 July 2021).

Gillingham, P. (2016) 'Predictive Risk Modelling to Prevent Child Maltreatment and Other Adverse Outcomes for Service Users: Inside the "Black Box" of Machine Learning', *The British Journal of Social Work*, 46(4), pp. 1044–1058. doi: 10.1093/bjsw/bcv031.

Gillingham, P. (2019) 'Decision Support Systems, Social Justice and Algorithmic Accountability in Social Work: A New Challenge', *Practice*, 31(4), pp. 277–290. doi: 10.1080/09503153.2019.1575954.

Gonzalez, M. G. (2020) *Cambridge City Council Votes to Ban Facial Recognition Technology*, *The Harvard Crimson*. Available at: https://www.thecrimson.com/article/2020/1/16/cambridge-city-council-bans-facial-recognition/ (Accessed: 7 July 2021).

Gonzalez, S. (2013) *In Newark, Gunshot Detection System Falls Short of Booker's Claims*, *WNYC*. Available at: https://www.wnyc.org/story/311533-gunshot-detection-sensors-newark-result-17-arrests-over-three-years/ (Accessed: 8 July 2021).

Government of Canada, O. of the A. G. of C. (2017) *Report 1—Phoenix Pay Problems*. Available at: https://www.oag-bvg.gc.ca/internet/English/parl_oag_201711_01_e_42666.html (Accessed: 8 July 2021).

Government of Canada, O. of the A. G. of C. (2018) *Report 1—Building and Implementing the Phoenix Pay System*. Available at: https://www.oag-bvg.gc.ca/internet/English/parl_oag_201805_01_e_43033.html (Accessed: 8 July 2021).

Gray, J. (2019) 'Toronto Police End ShotSpotter Project Over Legal Concerns', *The Globe and Mail*, 13 February. Available at: https://www.theglobeandmail.com/canada/toronto/article-toronto-police-end-shotspotter-project-over-legal-concerns/ (Accessed: 8 July 2021).

Guariglia, M. (2019) *Victory! Berkeley City Council Unanimously Votes to Ban Face Recognition*, *Electronic Frontier Foundation*. Available at: https://www.eff.org/deeplinks/2019/10/victory-berkeley-city-council-unanimously-votes-ban-face-recognition (Accessed: 7 July 2021).

Haag, D. (2020) *SyRI-wetgeving in strijd met het Europees Verdrag voor de Rechten voor de Mens*. Available at: https://www.rechtspraak.nl/Organisatie-en-contact/Organisatie/Rechtbanken/Rechtbank-Den-Haag/Nieuws/Paginas/SyRI-wetgeving-in-strijd-met-het-Europees-Verdrag-voor-de-Rechten-voor-de-Mens.aspx (Accessed: 18 August 2021).

Haskins, C. (2019) *Oakland Becomes Third U.S. City to Ban Facial Recognition*, *Vice*. Available at: https://www.vice.com/en/article/zmpaex/oakland-becomes-third-us-city-to-ban-facial-recognition-xz (Accessed: 7 July 2021).

Hegarty, P. (2019) *East Bay City Becomes Latest to Ban Use of Facial Recognition Technology*, *East Bay Times*. Available at: https://www.eastbaytimes.com/2019/12/18/east-bay-city-becomes-latest-to-ban-use-of-facial-recognition-technology (Accessed: 7 July 2021).

Hendry, J. (2018) *AFP Revives Investigation System Replacement*, *iTnews*. Available at: https://www.itnews.com.au/news/afp-revives-investigation-system-replacement-489890 (Accessed: 12 August 2021).

Henriques-Gomes, L. (2019) *Centrelink Still Issuing Incorrect Robodebts To Meet Targets, Staff Claim*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2019/may/29/centrelink-still-issuing-incorrect-robodebts-to-meet-targets-staff-claim (Accessed: 12 August 2021).

Henriques-Gomes, L. (2020) *Officials Refuse To Say How Many Flawed Robodebts They Have Found*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2020/mar/06/centrelink-robodebt-senate-estimates-class-action (Accessed: 12 August 2021).

Hern, A. (2020) *Ofqual's A-Level Algorithm: Why Did it Fail to Make the Grade?*, *the Guardian*. Available at: http://www.theguardian.com/education/2020/aug/21/ofqual-exams-algorithm-why-did-it-fail-make-grade-a-levels (Accessed: 15 July 2021).

Hesse, D. and Zürcher, C. (2018) 'Du wirst verbrochen haben', *Der Bund*, 11 May. Available at: https://www.derbund.ch/schweiz/standard/du-wirst-verbrochen-haben/story/13799968 (Accessed: 15 July 2021).

Heuberger, von S. (2017) *Predictive Policing: Software ist keine Wunderwaffe gegen Einbrüche*, *GQ*. Available at: https://www.gq-magazin.de/auto-technik/article/mit-big-data-den-einbrecher-schnappen-klappt-nur-bedingt (Accessed: 14 July 2021).

Hill, K. (2020) 'New Jersey Bars Police From Using Clearview Facial Recognition App', *The New York Times*, 25 January. Available at: https://www.nytimes.com/2020/01/24/technology/clearview-ai-new-jersey.html (Accessed: 8 July 2021).

Hoffmann, A. L. (2019) 'Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse', *Information, Communication & Society*, 22(7), pp. 900–915. doi: 10.1080/1369118X.2019.1573912.

Houston Federation of Teachers, Local 2415, et al., versus Houston Independent School District (2017).

Howgego, J. (2019) *A UK Police Force is Dropping Tricky Cases on Advice of an Algorithm*, *New Scientist*. Available at: https://www.newscientist.com/article/2189986-a-uk-police-force-is-dropping-tricky-cases-on-advice-of-an-algorithm/ (Accessed: 15 July 2021).

Human Rights Ombudsman (2018) *Annual Report of the Human Rights Ombudsman of the Republic of Slovenia for 2017*. 23rd. Ljubljana, Slovenia, p. 256.

Ito, J. (2018) 'Appendix B: Letter 1 to Massachusetts Legislature on Pretrial Risk Assessment', *Practice of Change*. Available at: https://www.practiceofchange.org/pub/n2wzxaft/release/1 (Accessed: 8 July 2021).

Jackson, D. and Marx, G. (2017) *Data Mining Program Designed to Predict Child Abuse Proves Unreliable, DCFS says*, *Chicago Tribune*. Available at: https://www.chicagotribune.com/investigations/ct-dcfs-eckerd-met-20171206-story.html (Accessed: 29 June 2021).

Jarmanning, A. (2020) *Boston Lawmakers Vote To Ban Use Of Facial Recognition Technology By The City*, NPR.org. Available at: https://www.npr.org/sections/live-updates-protests-for-racial-justice/2020/06/24/883107627/boston-lawmakers-vote-to-ban-use-of-facial-recognition-technology-by-the-city (Accessed: 7 July 2021).

JCWI (no date) *We Won! Home Office to Stop Using Racist Visa Algorithm*, *The Joint Council for the Welfare of Immigrants*. Available at: https://www.jcwi.org.uk/news/we-won-home-office-to-stop-using-racist-visa-algorithm (Accessed: 11 August 2021).

Karp, P. (2019a) *Government Admits Robodebt Was Unlawful As It Settles Legal Challenge*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2019/nov/27/government-admits-robodebt-was-unlawful-as-it-settles-legal-challenge (Accessed: 12 August 2021).

Karp, P. (2019b) *Robodebt: Government Abandons Key Part Of Debt Recovery Scheme In Major Overhaul*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2019/nov/19/robodebt-government-abandons-key-part-of-debt-recovery-scheme-in-major-overhaul (Accessed: 12 August 2021).

Keddell, E. (2015) 'The Ethics of Predictive Risk Modelling in the Aotearoa/New Zealand Child Welfare Context: Child Abuse Prevention or Neo-Liberal Tool?', *Critical Social Policy*, 35(1), pp. 69–88. doi: 10.1177/0261018314543224.

Kjær, J. S. (2018) *Regeringen har lagt sin plan om overvågning af børnefamilier i skuffen - politiken.dk*, *Politiken*. Available at: https://politiken.dk/indland/art6919255/Regeringen-har-lagt-sin-plan-om-overv%C3%A5gning-af-b%C3%B8rnefamilier-i-skuffen (Accessed: 14 July 2021).

Kolkman, D. (2020) '"F**k the Algorithm"?: What the World can Learn from the UK's A-Level Grading Fiasco', *Impact of Social Sciences*, 26 August. Available at: https://blogs.lse.ac.uk/impactofsocialsciences/2020/08/26/fk-the-algorithm-what-the-world-can-learn-from-the-uks-a-level-grading-fiasco/ (Accessed: 15 July 2021).

Lee, D. (2019) 'San Francisco is First US City to Ban Facial Recognition', *BBC News*, 14 May. Available at: https://www.bbc.com/news/technology-48276660 (Accessed: 7 July 2021).

'Letter to Los Angeles Police Commission' (2019). Available at: https://stoplapdspying.medium.com/on-tuesday-april-2nd-2019-twenty-eight-professors-and-forty-graduate-students-of-university-of-8ed7da1a8655 (Accessed: 7 July 2021).

Lipton, B. (2019) *Five Cities, Five Different Approaches to Algorithms in Government*, *MuckRock*. Available at: https://www.muckrock.com/news/archives/2019/jul/23/algorithms-five-cities/ (Accessed: 7 July 2021).

Mac, R., Haskins, C. and McDonald, L. (2020) *Clearview's Facial Recognition App Has Been Used By The Justice Department, ICE, Macy's, Walmart, And The NBA*, *BuzzFeed News*. Available at: https://www.buzzfeednews.com/article/ryanmac/clearview-ai-fbi-ice-global-law-enforcement (Accessed: 8 July 2021).

Macias Jr., M. (2019) *LAPD Says It Will Dump Data-Driven Policing Tools*, *Courthouse News Service*. Available at: https://www.courthousenews.com/lapd-says-it-will-dump-data-driven-policing-tools/ (Accessed: 7 July 2021).

Mackenzie, S. (2020a) *Police 'Stocktake' Surveillance Tech After Clearview AI Facial Recognition Trial*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/416913/police-stocktake-surveillance-tech-after-clearview-ai-facial-recognition-trial (Accessed: 13 August 2021).

Mackenzie, S. (2020b) *Police Trialled Facial Recognition Tech Without Clearance*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/416483/police-trialled-facial-recognition-tech-without-clearance (Accessed: 13 August 2021).

Markey, E. J. (2020) 'Letter to Mr. Hoan Ton-That'. US Senate.

Marsh, S. (2019) *One in Three Councils Using Algorithms to Make Welfare Decisions*, *the Guardian*. Available at: http://www.theguardian.com/society/2019/oct/15/councils-using-algorithms-make-welfare-decisions-benefits (Accessed: 15 July 2021).

Mayer, R. W. N. (2017) *Einbrüche in Stuttgart: Software soll Einbrecher stoppen*, stuttgarter-nachrichten.de. Available at: https://www.stuttgarter-nachrichten.de/inhalt.land-testet-einbruchs-software-kommissar-computer.ee56f-aed-3723-405d-8986-9aa76dfc888d.html (Accessed: 14 July 2021).

McDonald, H. (2020) *Home Office to Scrap 'Racist Algorithm' for UK Visa Applicants*, *the Guardian*. Available at: http://www.theguardian.com/uk-news/2020/aug/04/home-office-to-scrap-racist-algorithm-for-uk-visa-applicants (Accessed: 11 August 2021).

Medhora, S. (2019) *Federal Court Rules Robodebt 'Unlawful' in Significant Court Case*, *triple j*. Australian Broadcasting Corporation. Available at: https://www.abc.net.au/triplej/programs/hack/federal-government-loses-major-robodebt-case/11742494 (Accessed: 12 August 2021).

Ministry of Social Development (no date) *Vulnerable Children Predictive Modelling*, *Ministry of Social Development, New Zealand*. Available at: https://www.msd.govt.nz/about-msd-and-our-work/publications-resources/research/predicitve-modelling/index.html (Accessed: 13 August 2021).

Murphy, S. (2018) *Confirmation Overstayers Deported Using Profiling Programme*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/354270/confirmation-overstayers-deported-using-profiling-programme (Accessed: 12 August 2021).

Murray, A. and Giammarise, K. (2020) *Pittsburgh Suspends Policing Program that Used Algorithms to Predict Crime 'Hot Spots'*, *Pittsburgh Post-Gazette*. Available at: https://www.post-gazette.com/news/crime-courts/2020/06/23/Pittsburgh-suspends-policing-police-program-algorithms-predict-predictive-hot-spots-crime-data/stories/202006230059 (Accessed: 8 July 2021).

NCCPR Child Welfare Blog (2017) 'Los Angeles County Quietly Drops its First Child Welfare Predictive Analytics Experiment', *NCCPR Child Welfare Blog*, 9 May. Available at: https://www.nccprblog.org/2017/05/los-angeles-county-quietly-drops-its.html (Accessed: 29 June 2021).

NeoFace® Reveal: Advanced Criminal Investigation (no date) NEC. Available at: https://www.necam.com/AdvancedRecognitionSystems/Products/FacialRecognition/Solutions/NeoFaceReveal/ (Accessed: 8 July 2021).

Niklas, J. (2019) 'Poland: Government to Scrap Controversial Unemployment Scoring System', *AlgorithmWatch*, 16 April. Available at: https://algorithmwatch.org/en/poland-government-to-scrap-controversial-unemployment-scoring-system (Accessed: 15 July 2021).

Nikolic, I. (2020) 'French Court Rules against Facial Recognition in High Schools', *B.I.R.D.*, 28 February. Available at: https://bird.tools/french-court-rules-against-facial-recognition-in-high-schools/ (Accessed: 14 July 2021).

Nilsson, P. (2018) *First UK Police Force to Try Predictive Policing Ends Contract*, *Financial Times*. Available at: https://www.ft.com/content/b34b0b08-ef19-11e8-89c8-d36339d835c0 (Accessed: 15 July 2021).

Nilsson, P. (2019) 'UK Police Test if Computer can Predict Criminal Behaviour', *Financial Times*, 6 February. Available at: https://www.ft.com/content/9559efbe-2958-11e9-a5ab-ff8ef2b976c7 (Accessed: 15 July 2021).

North Tyneside Council (2019) *North Tyneside Council Report to Cabinet*.

North Tyneside Council Report to Cabinet Item 7(l) (2015).

North Tyneside Council, Revenue and Benefit Service (no date) *North Tyneside Risk Based Verification Policy*.

#NotMyDebt (2016) *The Issue*. Not My Debt. Available at: https://www.notmydebt.com.au/the-issue (Accessed: 12 August 2021).

n-tv.de (2019) *Gewerkschaft kann Aus* für *Precobs-Software nachvollziehen*, n-tv.de. Available at: https://www.n-tv.de/regionales/baden-wuerttemberg/Gewerkschaft-kann-Aus-fuer-Precobs-Software-nachvollziehen-article21250215.html (Accessed: 14 July 2021).

Office of the Auditor General (2016) *Michigan IntegratedData Automated System (MiDAS)*. Performance Audit Report 641-0593–15.

Office of the Inspector General (2019) *Review of Selected Los Angeles Police Department Data-Driven Policing Strategies*. Los Angeles: Los Angeles Police Commission, p. 51.

Office of the Privacy Commissioner of Canada (2020a) *Clearview AI Ceases Offering its Facial Recognition Technology in Canada*, *Office of the Privacy Commissioner of Canada*. Available at: https://www.priv.gc.ca/en/opc-news/news-and-announcements/2020/nr-c_200706/ (Accessed: 8 July 2021).

Office of the Privacy Commissioner of Canada (2020b) *Commissioners Launch Joint Investigation Into Clearview AI Amid Growing Concerns Over Use Of Facial Recognition Technology*, *Office of the Privacy Commissioner of Canada*. Available at: https://priv.gc.ca/en/opc-news/news-and-announcements/2020/an_200221/ (Accessed: 8 July 2021).

Office of the Privacy Commissioner of Canada (2020c) *OPC Launches Investigation Into RCMP's Use of Facial Recognition Technology*, *Office of the Privacy Commissioner of Canada*. Available at: https://priv.gc.ca/en/opc-news/news-and-announcements/2020/an_200228/ (Accessed: 8 July 2021).

Oswald, M. *et al.* (2017) *Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and 'Experimental' Proportionality*. SSRN Scholarly Paper ID 3029345. Rochester: Social Science Research Network. doi: 10.2139/ssrn.3029345.

Pascu, L. (2020) *French High Court Rules Against Biometric Facial Recognition Use in High Schools*, *Biometric Update*. Available at: https://www.biometricupdate.com/202002/french-high-court-rules-against-biometric-facial-recognition-use-in-high-schools (Accessed: 14 July 2021).

Patton, J. (2018) *What Is 'ShotSpotter'? Controversial Gunshot Detector Technology Approved By Toronto Police*, Globalnews.ca. Available at: https://globalnews.ca/news/4344093/controversial-gunshot-detector-shotspotter-toronto-police/ (Accessed: 8 July 2021).

Pennington, P. (2020a) *Government Facial Recognition Tech Deal Offers Wide Access*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/428103/government-facial-recognition-tech-deal-offers-wide-access (Accessed: 13 August 2021).

Pennington, P. (2020b) *Police Setting Up $9m Facial Recognition System Which Can Identify People From CCTV Feed*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/424845/police-setting-up-9m-facial-recognition-system-which-can-identify-people-from-cctv-feed (Accessed: 13 August 2021).

Pett, H. and Cosier, C. (2017) 'We're All Talking About Centrelink, But What Is "Robodebt" Anyway?', *ABC News*, 2 March. Available at: https://www.abc.net.au/news/2017-03-03/centrelink-debt-controversy-what-is-robodebt/8317764 (Accessed: 12 August 2021).

Pitt Cyber (2020) *Pittsburgh Task Force on Public Algorithms*, *Institute for Cyber Law, Policy, and Security*. Available at: https://www.cyber.pitt.edu/algorithms (Accessed: 8 July 2021).

Porter, J. (2020) *UK Ditches Exam Results Generated by Biased Algorithm After Student Protests*, *The Verge*. Available at: https://www.theverge.com/2020/8/17/21372045/uk-a-level-results-algorithm-biased-coronavirus-covid-19-pandemic-university-applications (Accessed: 15 July 2021).

Predictive Policing Evaluation of the Pilot Project P4 in Baden-Württemberg/Germany (2018) *Max Planck Institute for the Study of Crime, Security and Law*. Available at: http://link.springer.com/10.1007/s41125-018-0033-0 (Accessed: 14 July 2021).

Privacy Commissioner, New Zealand (2015) *Predicting Risk Without Turning Children Into 'Lab Rats'*, *Privacy Commissioner*. Available at: https://www.privacy.org.nz/blog/predictive-risk-modelling/ (Accessed: 13 August 2021).

Privacy International (2020) *The SyRI Case: A Landmark Ruling for Benefits Claimants Around the World*, *Privacy International*. Available at: http://privacyinternational.org/news-analysis/3363/syri-case-landmark-ruling-benefits-claimants-around-world (Accessed: 14 July 2021).

Robertson, K., Khoo, C. and Song, Y. (2020) *To Surveil and Predict: A Human Rights Analysis of Algorithmic Policing in Canada*, *The Citizen Lab*. Available at: https://citizenlab.ca/2020/09/to-surveil-and-predict-a-human-rights-analysis-of-algorithmic-policing-in-canada/ (Accessed: 8 July 2021).

Scharfenberg, D. (2018) *Computers can Solve Your Problem. You May Not Like the Answer.*, BostonGlobe.com. Available at: https://apps.bostonglobe.com/ideas/graphics/2018/09/equity-machine (Accessed: 7 July 2021).

Senate of Canada (2018) *The Phoenix Pay Problem: Working Toward a Solution*. Ottawa: Standing Senate Committee on National Finance. Available at: https://sencanada.ca/en/info-page/parl-42-1/nffn-phoenix-pay-problem/ (Accessed: 8 July 2021).

Settlement and Full and Final Release Agreement (2017). State of Texas County of Harris.

Sheridan, E. (2019) 'Town Hall Drops Pilot Programme Profiling Families Without Their Knowledge', *Hackney Citizen*, 30 October. Available at: https://www.hackneycitizen.co.uk/2019/10/30/town-hall-drops-pilot-programme-profiling-families-without-their-knowledge/ (Accessed: 15 July 2021).

Simonite, T. (2020) *Algorithms Were Supposed to Fix the Bail System. They Haven't*, *Wired*. Available at: https://www.wired.com/story/algorithms-supposed-fix-bail-system-they-havent/ (Accessed: 8 July 2021).

SRF (2017) *Solothurner Polizei will keine Software zur Einbruchsbekämpfung*, *Schweizer Radio und Fernsehen (SRF)*. Available at: https://www.srf.ch/news/regional/aargau-solothurn/aargauer-solothurner-polizei-solothurner-polizei-will-keine-software-zur-einbruchsbekaempfung (Accessed: 15 July 2021).

STA (2019) *Automatic Licence Plate Recognition Banned in Slovenia*, *Total Slovenia News*. Available at: https://www.total-slovenia-news.com/politics/4126-automatic-licence-plate-recognition-banned-in-slovenia (Accessed: 15 July 2021).

STA (2020) *Interview: Police Chief on Cybercrime, Dark Web, Drugs & Illegal Migration*, *Total Slovenia News*. Available at: https://www.total-slovenia-news.com/politics/5291-interview-police-chief-on-cybercrime-dark-web-drugs-illegal-migration (Accessed: 15 July 2021).

Statewatch (2020) *UK: Legal Action Threatened over Algorithm Used to Grade Teenagers' Exams*, *Statewatch*. Available at: https://www.statewatch.org/news/2020/august/uk-legal-action-threatened-over-algorithm-used-to-grade-teenagers-exams/ (Accessed: 15 July 2021).

Stone, L. (2020) *UK Police Scrap AI Prototype Designed to Predict Serious Violence*, *AIBotics*. Available at: https://www.aibotics.tech/post/uk-police-scrap-ai-prototype-designed-to-predict-serious-violence (Accessed: 12 August 2021).

Stop LAPD Spying Coalition (2018) *Before the Bullet Hits the Body – Dismantling Predictive Policing in Los Angeles*. Los Angeles: Stop LAPD Spying Coalition. Available at: https://stoplapdspying.org/before-the-bullet-hits-the-body-dismantling-predictive-policing-in-los-angeles/ (Accessed: 7 July 2021).

Tay, L. (2011) *Integration Issues Halt Vic Police's LINK Project*, *iTnews*. Available at: https://www.itnews.com.au/news/integration-issues-halt-vic-polices-link-project-255133 (Accessed: 12 August 2021).

tellerreport.com (2019) *Belgian Police Stop Facial Recognition At Zaventem Airport*, Tellerreport.com. Available at: https://www.tellerreport.com/tech/2019-09-21---belgian-police-stop-facial-recognition-at-zaventem-airport-.BkEeQN8QDH.html (Accessed: 8 July 2021).

Tunney, C. (2020) *RCMP Says It Will Limit Its Use of Facial Recognition Tech — But Won't Stop Using It Entirely*, *CBC*. Available at: https://www.cbc.ca/news/politics/rcmp-clearview-ai-1.5490988 (Accessed: 8 July 2021).

United Nations Special Rapporteur (2020) 'Brief by the United Nations Special Rapporteur on extreme poverty and human rights as Amicus Curiae in the case of NJCM C.s./De Staat der Nederlanden (SyRI) before the District Court of The Hague (case number: C/09/550982/HA ZA 18/388)'.

University of Cambridge (2018) *Helping Police Make Custody Decisions Using Artificial Intelligence*, *University of Cambridge*. Available at: https://www.cam.ac.uk/research/features/helping-police-make-custody-decisions-using-artificial-intelligence (Accessed: 15 July 2021).

University of Copenhagen (2019) *New Technology to Prevent Future IT Flops in the Public Sector*, *Faculty of Science, University of Copenhagen*. University of Copenhagen. Available at: https://science.ku.dk/english/press/news/2019/new-technology-to-prevent-future-it-flops-in-the-public-sector/ (Accessed: 14 July 2021).

Victorian Government Health Information (2012) *HealthSMART - Victorian Health IT*. Available at: https://web.archive.org/web/20120618013645/http:/www.health.vic.gov.au/healthsmart/ (Accessed: 12 August 2021).

Wakefield, J. (2020) 'A-Levels: Ofqual's "Cheating" Algorithm Under Review', *BBC News*, 20 August. Available at: https://www.bbc.com/news/technology-53836453 (Accessed: 15 July 2021).

Webb, S. (2017) *Houston Teachers to Pursue Lawsuit over Secret Evaluation System*, *Houston Chronicle*. Available at: https://www.houstonchronicle.com/news/houston-texas/houston/article/Houston-teachers-to-pursue-lawsuit-over-secret-11139692.php (Accessed: 29 June 2021).

Why Do Projects Fail? (2012) 'HealthSMART – Victoria', *Why Do Projects Fail?*, 11 November. Available at: http://calleam.com/WTPF/?p=3847 (Accessed: 12 August 2021).

Wikipedia (2021) 'PRECOBS', *Wikipedia*. Available at: https://de.wikipedia.org/w/index.php?title=PRECOBS&oldid=210617276 (Accessed: 15 July 2021).

Wills, T. (2019) 'Sweden: Rogue Algorithm Stops Welfare Payments for up to 70,000 Unemployed', *AlgorithmWatch*, 25 February. Available at: https://algorithmwatch.org/en/rogue-algorithm-in-sweden-stops-welfare-payments (Accessed: 15 July 2021).

Wisterberg, E. (2018a) *12 av 16 säger upp sig i protest mot kommunens bidragsrobot*, *Breakit*. Available at: https://www.breakit.se/artikel/11082 (Accessed: 15 July 2021).

Wisterberg, E. (2018b) *Titta! Kommunens robot avgör vem som får socialbidrag…*, *Breakit*. Available at: https://www.breakit.se/artikel/10770 (Accessed: 15 July 2021).

Yogaretnam, S. (2020) *Chief Says Facial Recognition Software Test Drew on Police Mug Shot Database*, *Ottawa Citizen*. Available at: https://ottawacitizen.com/news/local-news/chief-says-facial-recognition-software-test-drew-on-police-mug-shot-database (Accessed: 8 July 2021).

collective wellbeing Carnegie UK

Data Justice Lab

Western FIMS

Automating Public Services: Learning from Cancelled Systems

Case Studies

Joanna Redden, Jessica Brand, Ina Sander and Harry Warne
Data Justice Lab

Contents

Introduction

These case study overviews, which draw upon interviews and document analysis, provide: a summary, details about why the system was cancelled, background information, key factors identified as leading to change, information about how the system worked, details about positive or negative impacts as well as what is known about the case going forward.

Fraud Detection ADS

Overview

Our case studies demonstrate the need to be alert to how the use of ADS creates differential systems of advantage and disadvantage (Crenshaw, 1989; Hoffmann, 2019). We found a number of similarities across countries in their use of automated benefit fraud detection systems. The systems researched as part of our case studies have been publicly criticized such as: Netherlands System Risk Indication (SyRI); the Australian Online Compliance Intervention system (RoboDebt), the Michigan Integrated Data Automated System and the use of Risk Based Verification Systems in the UK.

Case Studies: Fraud Detection

Netherlands	Ministry of Social Affairs and Employment System Risk Indication (SyRI) (2014-2020)
United States, Michigan	Michigan Unemployment Insurance Agency stops using Michigan Integrated Data Automated System (Midas) for automated fraud assessments (2013-2015)
Australia	Robo-debt / Online Compliance Intervention Stopped (2016 – 2019)
UK	Several local authorities stop using automated risk-based verification systems (2013 - 2020).

Presumption of guilt

One of the elements of criticism that unites these systems is the presumption of guilt being attached to those applying for or using benefits. The point was raised by former UN Special Rapporteur on Extreme Poverty and Human Rights, Philip Alston, in his report to the UK (2019).

'The presumption of innocence is turned on its head when everyone applying for a benefit is screened for potential wrongdoing in a system of total surveillance. And in the absence of transparency about the existence and workings of automated systems, the rights to contest an adverse decision, and to seek a meaningful remedy, are illusory.' (Alston, 2018)

The concern being raised is how people with low incomes who are applying for benefits are differentially targeted by surveillance and sorting systems. Key in all of the above listed systems, with the exception of the UK's RBV systems, is that fraud notices were automatically generated with those flagged sent letters telling them they needed to prove their innocence. Guilt in these cases is assumed, rather than it being the responsibility of the state to prove guilt. It is difficult to imagine a similar system being applied to tackle white collar crime, but this is an important thought exercise. If a system has been put in place to automatically generate fraud or overpayment notices with the onus then placed on the recipient to prove innocence, why was a similar system not put in place in these countries to automatically generate notices of tax evasion, money laundering or insider trading with onus on the recipient to prove innocence.

Opaqueness and inflexibility

It is deeply problematic to introduce systems to automatically assess overpayment of benefits and then make it difficult for people to get access to information about the system's decisions. In the case of the Robodebt scandal in Australia, people reported spending hours trying to speak with a compliance officer and then once connected being told they could not have access to the information about their case. The problems with this are intensified when it is recognized that error will always occur with automated decision systems, as there is no perfect system. Further, the lack of information available once automatically flagged makes it difficult to challenge the system and in some cases for people to know that the process is automated.

Justice at what cost?

The legality of some of these systems have been successfully challenged but doing so has taken years. In the meantime, reports detail the harm caused. As detailed in our case studies, errors have led to wrongful debt collection, persecution, bankruptcy, stress, family breakdown and illness. In many cases, when the error is finally accepted, and the money returned it is too late as the harm has been done. In the meantime, much community resources, time, energy, and money has been used to challenge the systems. In the Netherlands, SyRI was in place for over six years, and in 2020 a court ruled the system violates human rights. Australia's robo-debt system was in place for three years and was only ruled unlawful in 2020 with a class action lawsuit settled for $1.2 billion. In the United States, the Automated Food Assistance Eligibility system and Michigan's Integrated data Automated System were successfully legally challenged. Other systems were cancelled by government agencies after years in use because they were determined not to be effective.

Netherlands, SyRI

Summary

The Systeem Risico Inventarisatie (System Risk Indication or SyRI) is a big data analysis system that ran under the Ministry of Social Affairs and Employment, aiming to assess risks of welfare abuse and tax fraud. The nationwide system had been in use since 2008, and was passed into law in 2014. It was abandoned in April 2020 after it was found to be unlawful and not compliant with Article 8 paragraph 2 of the European Convention on Human Rights released on Feb. 5, 2020 (European Court of Human Rights, 2020).

Why was it stopped?

After a coalition of privacy and civil rights groups brought a lawsuit against the Dutch government's use of SyRI, the District Court of the Hague ruled that the legislation governing the deployment of SyRI violates higher law and does not comply with Article 8 of the European Convention on Human Rights (ECHR). This provision requires a fair balance (a reasonable relationship) between the social interest served by the legislation and the infringement of private life.

In response, the State Secretary for Social Affairs and Employment, Tamara van Ark (van Ark, 2020), said she had decided not to appeal this decision because the system was not efficient or effective and was instead going to investigate how to improve the technologies for fraud detection. Media reports also suggest that the system did not identify new cases, only previously identified ones.

Key factors leading to change

Civil Society Mobilization: This mobilization involved strategic litigation, coalition building, and a public awareness campaign which included publications and meetings with residents. One of the lawyers involved with the case thinks mobilizing public opinion was key to winning the case.

The lawsuit against the Dutch State was initiated by a coalition of civil society organisations, consisting of the Platform for the Protection of Civil Rights (Platform Bescherming Burgerrechten), the Dutch Section of the International Commission of Jurists (NJCM), trade union FNV, Privacy First, Foundation KDVP, the National Client Council and authors Tommy Wieringa and Maxim Februari. PILP-NJCM coordinated the court case. The case was handled by lawyers Anton Ekker of Ekker Advocatuur and Douwe Linders of SOLV." (https://pilpnjcm.nl/en/dossiers/profiling-and-syri/)

How did the system work?

Government representatives say the system was introduced to fight fraud to "maintain social security support." Civil society advocates think it was introduced to save money and to signal a tough on crime approach.

The Systeem Risico Inventarisatie (System Risk Indication or SyRI) was a big data analysis system run by the Ministry of Social Affairs and Employment. State institutions could use the system to try and detect those trying to unlawfully access public money or benefits. The system combines data and an algorithmic model is used to analyse these combined data sets in order to detect irregularities or potential fraud (Algorithm Watch, 2019).

The types of data linked and analysed included personal data such as identity data, data related to labour, property, education, pension, business, income, assets, and debts.

The system produced risk reports on addresses identified as presenting an increased risk of fraud. People were registered and could be investigated. One of the main criticisms of the system, and put forward by the Public Interest Litigation Project, is that everyone in the Netherlands was "suspected in advance"[1] through the government's use of this system.

The system was used at a local level by collaborating public authorities and was used nationwide. Algorithm Watch reported that the system has been used in Capelle aan den Ijssel, Eindhoven, Haarlem and Rotterdam.[2]

Case notes contain a list of datasets that could be used as well as the process for accessing the data (Haag, 2020). The risk model looked for risk indicators which considered "behavioural norms." The judge was critical of the lack of transparency about the system, that it was difficult to find out what the system did and how (Haag, 2020).

Concerns were raised by civil rights advocates that the system does on respect the data protection principle of limitation, that citizens are not informed of the software and of the scores they may get or who has access to this information.

Concerns were also raised about which neighbourhoods were targeted by system, particularly that the system was used in areas with high numbers of marginalized populations (Hendrickx Interview, 2020).

There were efforts to notify people in some of these communities about the system, which led to civil society action. It has been argued that this public response, in part, led the Mayor of Rotterdam to stop using the system (Hendrickx Interview, 2020).

Researchers assessing the data protection and privacy implications of SyRI raised concerns about risk reports being recorded in SyRI for two years.[3] The Dutch government began experimenting with large scale algorithmic risk assessment in 2007 and developed risk models and risk indicators for years before SyRI was adopted and before legislation to enable SyRI was introduced. Some of these cases involved examining entire communities by the system.

Government officials noted that data was pseudonymized when given to the trusted third party to run the analysis and that only "hits" would be deanonymized. Civil society actors argue that combining and using data on every citizen renders every citizen a suspect (Tijmen Interview, 2020).

There was a data protection impact assessment when the law was introduced, but it seems impact assessments were not done in relation to specific projects.

The system involved corporate partnership as different parties provided their data to the Foundation Intelligence Agency and a private foundation Stichting Inlichtingenbureau (the Benefits Intelligence Agency Foundation) processes the data to determine risk.

1 See statement by PILP here: https://pilpnjcm.nl/en/dossiers/profiling-and-syri/

2 Ilja Braun (2018) 'High Risk Citizens,' Algorithm Watch, 4 July 2018, available: https://algorithmwatch.org/en/high-risk-citizens/

3 van Dalen, Steven; Gilder, Alexander; Hooydonk, Eric; Ponsen, Marc (2016) System Risk Indication: An Assessment of hte Dutch Anti-Fraud System in the Context of Data Protection and Profiling. Public Interest Litigation Project, available: https://pilpnjcm.nl/wp-content/uploads/2016/06/memorandum_1_-_system_risk_indication_an_assessment_of_the_dutch_anti-fraud_system_in_the_context_of_data_protection_and_profiling-1.pdf

Impact

In the case of Capelle aan den Ijssel it was reported that the use of SyRI led to the identification of violations and recovery of benefits, although it is not clear if these identifications were new cases.

Initially, privacy and human rights organisations opposed the system. There was critical newspaper coverage. Local constitutional councils like the data protection supervisor were critical. One interviewee noted that a famous essayist, Tommy Wieringa, wrote a spoken word essay that got a lot of attention and captured people's interest (Hussein Interview, 2020). Other famous writers also got involved. A Dutch trade union, for instance, started mobilizing people in communities.

When the trade union FNV got involved, they started holding public meetings in the two communities in Rotterdam they knew were being targeted by Syri investigations. People in those communities voiced concerns about bias, wondering why they were being targeted. There was also worry by residents that they could be on a watch list and not know, particularly since this could be the product of unintentional errors in filling in tax information or other kinds of forms. Local politicians got involved and asked questions of the Mayor of Rotterdam. The mayor stopped using the system, which got a lot of public attention. It was suggested that this helped change public opinion (Hussein Interview, 2020).

The former UN Special Rapporteur on Extreme Poverty and Human Rights, Philip Alston, wrote an Amicus Curiae for the court case which also sparked media attention. He explains his two key concerns: 1. The "human right to social security" (Alston, 2018, par. 19ff) and "the right to privacy" (Alston, 2018, par. 28ff). The Digital Freedom Foundation appeared in court.

Key ideas to emerge

Use of the system meant that while it should be that everyone is entitled to a right of privacy, those who live in addresses targeted by SyRI do not have this right (Wisman Interview 2020, citing Alston).

One of the reasons for the court case was to challenge the idea that everyone is a suspect, that with a system like this you are no longer innocent until proven guilty but a suspect until proven innocent (Wisman Interview, 2020).

The judge argued that more transparency was needed so the system could be checked and people could challenge the system and the data.

The ruling challenges the view that a risk notification system is just data, as it presents a legal position that data held about people can have an impact (Hussein and Wisman Interviews 2020).

Going forward

Public authorities in the Netherlands announced that they would revise their fraud systems in light of this decision. The Dutch Employment Agency is reviewing its internal fraud systems. The Dutch Tax Service ceased the operation of a fraud detection system after facing Dutch Data Protection Authority investigation.

In a letter to the President of the House of Representatives, State Secretary for Social Affairs and Employment, Tamara van Ark said she is interested in developing a new instrument and in studying how "new technology can be used to combat fraud effectively and efficiently, with adequate safeguards for privacy."

Sources

Interviews conducted online in 2020:

Tijmen Wisman: "Chairman of the Platform for Civil Rights and Assistant Professor of privacy law at the Vrije Universiteit of Amsterdam," involved in legal case

Ronald Huissen: Secretary for Platform for Civil Rights, involved in political and public lobbying related to the case, research for case

Merel Hendrickx: In-house human rights Lawyer with the Public Interest Litigation Project of the Dutch Section of the International Commission of Jurists (PILP-NJCM)."

Interview requests sent to two government officials were declined.

References

Algorithm Watch (2019) 'Automating Society', *AlgorithmWatch*. Available at: https://algorithmwatch.org/en/automating-society/ (Accessed: 18 August 2021).

Alston, P. (2018) *Statement on Visit to the United Kingdom, by Professor Philip Alston, United Nations Special Rapporteur on extreme poverty and human rights*, *United Nations Human Rights Office of the High Commissioner*. Available at: https://www.ohchr.org/en/NewsEvents/Pages/DisplayNews.aspx?NewsID=23881 (Accessed: 18 August 2021).

van Ark, T. (2020) 'Letter of 23 April 2020 to the President of the House of Representatives by the State Secretary for Social Affairs and Employment, Tamara van Ark, on a Court Judgment Regarding SyRI'.

Bij Voorbaat Verdacht (2019) 'Datum zitting bodemprocedure tegen SyRI op 29 oktober 2019', *Bij Voorbaat Verdacht*, 23 April. Available at: https://bijvoorbaatverdacht.nl/save-the-date-zitting-rechtszaak-tegen-syri-op-29-oktober-2019/ (Accessed: 19 August 2021)

Bij Voorbaat Verdacht (2020a) 'Inbreng Platform Burgerrechten aan de Tweede Kamer over "Super SyRI"', *Bij Voorbaat Verdacht*, 26 May. Available at: https://bijvoorbaatverdacht.nl/lees-onze-input-aan-de-tweede-kamer-over-super-syri/ (Accessed: 19 August 2021).

Bij Voorbaat Verdacht (2020b) 'Kamer stuurt aan op grondige behandeling Super SyRI', *Bij Voorbaat Verdacht*, 5 June. Available at: https://bijvoorbaatverdacht.nl/kamer-stuurt-aan-op-grondige-behandeling-super-syri/ (Accessed: 19 August 2021).

Bij Voorbaat Verdacht (2020c) 'Super SyRI: de volgende stap naar de kijkdoosmaatschappij', *Bij Voorbaat Verdacht*, 25 May. Available at: https://bijvoorbaatverdacht.nl/super-syri-de-volgende-stap-naar-de-kijkdoosmaatschappij/ (Accessed: 19 August 2021).

Bij Voorbaat Verdacht (2020d) 'SyRI-coalitie maant kabinet: stop overhaaste invoering "Super SyRI"', *Bij Voorbaat Verdacht*, 25 May. Available at: https://bijvoorbaatverdacht.nl/syri-coalitie-maant-kabinet-stop-overhaaste-invoering-super-syri/ (Accessed: 19 August 2021).

Blauw, S. (2020a) *An Algorithm Was Taken To Court – And It Lost (Which Is Great News For The Welfare State)*, *The Correspondent*. Available at: https://thecorrespondent.com/276/an-algorithm-was-taken-to-court-and-it-lost-which-is-great-news-for-the-welfare-state (Accessed: 19 August 2021).

Blauw, S. (2020b) *How a Small Group of Activists Put Big Brother on Trial – And Won*, *The Correspondent*. Available at: https://

thecorrespondent.com/344/how-a-small-group-of-activists-put-big-brother-on-trial-and-won (Accessed: 19 August 2021).

Braun, I. (2018) 'High-Risk Citizens', *AlgorithmWatch*, 4 July. Available at: https://algorithmwatch.org/en/high-risk-citizens/ (Accessed: 19 August 2021).

Crenshaw, K. (2015) 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', *University of Chicago Legal Forum*, 1989(1). Available at: https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8

Deikwijs Advocaten (no date) 'Subpeona'.

Digital Freedom Fund (2020) 'NJCM, Platform Bescherming Burgerrechten and others v. The Netherlands (the SyRI case)', *Digital Freedom Fund*, February. Available at: https://digitalfreedomfund.org/case-analyses/njcm-platform-bescherming-burgerrechten-and-others-v-the-netherlands/ (Accessed: 19 August 2021).

European Court of Human Rights (2020) *European Convention on Human Rights*. France: Council of Europe, p. 34.

Gantchev, V. (2019) 'Data Protection in The Age of Welfare Conditionality: Respect For Basic Rights or a Race to The Bottom?', *European Journal of Social Security*, 21(1), pp. 3–22. doi: 10.1177/1388262719838109.

Haag, D. (2020) *SyRI-wetgeving in strijd met het Europees Verdrag voor de Rechten voor de Mens*, *de Rechtspraak*. Available at: https://www.rechtspraak.nl/Organisatie-en-contact/Organisatie/Rechtbanken/Rechtbank-Den-Haag/Nieuws/Paginas/SyRI-wetgeving-in-strijd-met-het-Europees-Verdrag-voor-de-Rechten-voor-de-Mens.aspx (Accessed: 18 August 2021).

Henley, J. and Booth, R. (2020) *Welfare Surveillance System Violates Human Rights, Dutch Court Rules*, *the Guardian*. Available at: http://www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules (Accessed: 19 August 2021).

Hoffmann, A. L. (2019) 'Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse', *Information, Communication & Society*, 22(7), pp. 900–915. doi: 10.1080/1369118X.2019.1573912.

Jenezon, J. and Smits, J. (2019) *SyRI: Systeem Risico-indicatie*. ministerie van Sociale Zaken en Werkgelegenheid.

Minister voor Rechtsbescherming (2018) *Wet gegevensverwerking door samenwerkingsverbanden*. Available at: https://wetgevingskalender.overheid.nl/Regeling/WGK008727 (Accessed: 19 August 2021).

Nederlands Juristen Comité voor de Mensenrechten (2018) *Rights Groups Sue Dutch State for Profiling Citizens*, *Liberties.eu*. Available at: https://www.liberties.eu/en/stories/ngos-start-proceedings-against-dutch-government-over-syri/14251 (Accessed: 19 August 2021).

Privacy First (2018) *Civil Rights Groups Sue Dutch Government For Risk Profiling Citizens*, *Privacy First*. Available at: https://www.privacyfirst.eu/court-cases/667-civil-rights-groups-sue-dutch-government-for-risk-profiling-citizens.html (Accessed: 19 August 2021).

Privacy International (2020) *The SyRI Case: A Landmark Ruling For Benefits Claimants Around The World*, *Privacy International*. Available at: http://privacyinternational.org/news-analysis/3363/syri-case-landmark-ruling-benefits-claimants-around-world (Accessed: 19 August 2021).

The Hague District Court (2020) *ECLI:NL:RBDHA:2020:1878, Rechtbank Den Haag, C-09-550982-HA ZA 18-388 (English)*. Available at:

https://deeplink.rechtspraak.nl/uitspraak?id=ECLI:NL:RBDHA:2020:1878 (Accessed: 19 August 2021).

The Public Interest Litigation Project (2015) *Profiling and SyRI*, *PILP*. Available at: https://pilpnjcm.nl/en/dossiers/profiling-and-syri/ (Accessed: 19 August 2021).

Tweede Kamer der Staten-Generaal (2013) *Wijziging van de Wet structuur uitvoeringsorganisatie werk en inkomen en enige andere wetten in verband met fraudeaanpak door gegevensuitwisselingen en het effectief gebruik van binnen de overheid bekende zijnde gegevens; Memorie van toelichting; Memorie van toelichting*. officiële publicatie 33579 nr. 3. Available at: https://zoek.officielebekendmakingen.nl/kst-33579-3.html (Accessed: 19 August 2021).

United Nations Special Rapporteur (2020) 'Brief by the United Nations Special Rapporteur on extreme poverty and human rights as Amicus Curiae in the case of NJCM C.s./De Staat der Nederlanden (SyRI) before the District Court of The Hague (case number: C/09/550982/HA ZA 18/388)'.

Wisman, T. (2020) 'The SyRI Victory: Holding Profiling Practices to Account', *Digital Freedom Fund*, 23 April. Available at: https://digitalfreedomfund.org/the-syri-victory-holding-government-profiling-to-account/ (Accessed: 19 August 2021).

Michigan Unemployment Insurance Agency stops using Michigan Integrated Data Automated System (Midas) for automated fraud assessments

Summary

In 2013 Michigan's Unemployment Insurance Agency (UIA) launched MiDAS to administer and process unemployment benefits. A key component of the system was its use for automated fraud assessment. It has been reported that the number of people suspected of fraud grew drastically after the system was implemented and that within two years the MiDAS system generated false accusations of fraud for 40,000 people (de la Garza 2020). There have been numerous court proceedings related to people's experiences.

Why was it stopped?

It has been argued that UIA stopped using MiDAS for automated fraud assessment in September 2015 due to federal government pressure and a federal lawsuit (Charette 2018; Fleming and Fournier 2015). The UIA apologised in January 2017. In 2017 the Michigan State legislature passed a law requiring that fraud detection be done manually (de la Garza, 2020). Legal proceedings continue to try and have people remunerated for the damage caused to their lives due to charges of fraud and financial penalties introduced in error. Media coverage has detailed a range of harms caused by the system including bankruptcy, eviction, and homelessness (de la Garza, 2020; Egan, 2019; Charette 2018).

Key factors leading to change

There has been ongoing widespread citizen and community level push back against the use of this system.

The Michigan Unemployment Insurance Clinic and other community organisations were contacted by those who were wrongly accused of fraud. Members of these organisations sent a letter to the US Department of Labour expressing deep concern and urging investigation. There was also critical media coverage (Felton, 2015), legal action and two audits of the UIA. Michigan Auditor General reports provided key details, including finding 40,000 false fraud cases, a 93 percent inaccuracy rate and that 96 percent of people making calls to the Agency were being ignored.

Michigan's state representative at the time, Sandy Levin, wrote a damning public letter (Levin, 2016) in April 2016 to Governor Snyder urging him to "review the claims of fraud made by the State Unemployment Insurance Agency's automated system and fully reimburse those citizens of our state who were harmed due to inaccurate determinations" (Levin, 2016).

In addition, lawsuits have been filed against the Agency. While a number were dismissed, Zynda et al v Zimmer et al has been cited as a key case that resulted in a settlement that ordered the UIA to cease automating adjudications and introduce human oversight (*2:15-cv-11449-RHC-RSW*, 2017). According to one of the Attorneys who represented the plaintiffs on the Zynda case, MiDAS "criminalizes unemployment" (Felton, 2015). Lord noted that a group of lawyers came together and took that case to federal court and were able to secure changes.

Lord argues that that a basic claim of the lawsuit she has been involved in is that this automated system meant that the State was accusing people and taking their money without due process:

"[T]he premise that someone wouldn't get compensation if their life or their liberty or their property were taken without due process is highly offensive to me, and I'm beyond baffled why the government would be advocating for that position against its own citizens" (Lord interview, 2020).

It's been argued that the public and political outcry in combination with legal actions led the UIA to admit there was a problem with its 'robo-adjudication' process (Charette 2018).

Sandy Levin's letter suggests that pressure from the Department of Labor was what prompted the Agency to stop automating MiDAS: he writes "I understand that because of oversight efforts by the U.S. DOL (DOL), the computer system is no longer in violation of federal law and has placed an employee back in charge of verifying any fraud allegation" (Levin, 2016). Ryan Felton's in-depth news coverage shows that, when the Zynda federal lawsuit was filed in April 2015, the state had until mid-August 2015 to respond to the lawsuit (Felton, 2015), which possibly suggests that the threat of legal action was enough for the UIA to suspend automation.

How did the system work?

The system involved funding through the REED Act and the American Recovery and Reinvestment Act (ARRA). The UIA received up to 69.5 million for financing the MiDAS project of up to $69.5 million (McFarlane, 2014: p.4).

Three private vendors were contracted to help with the development MiDAS: SAS Institute, FAST Enterprises LLC and CSG Government Solutions. FAST Enterprises was the main developer of the system and in July 2011 the state awarded Fast Enterprises the multi-million dollar contract to begin work on the MiDAS project, "a deal officials pegged as a modernization of the state›s unemployment insurance infrastructure" (Felton 2015). The contract to build the system was for more than $47 million1 (Wykstra, 2020) and took 26 months to develop (Charette, 2018).

The UIA awarded an $18 million bid to Chicago-based CSG Government Solutions to provide a "full-time, onsite project management [team] to oversee a comprehensive and complex rewrite of Michigan's current Unemployment Insurance Systems." The state also awarded a $14 million contract ('Contract Change Notice', 2018) to SAS Analytics to implement fraud detection software that would be integrated into MiDAS. The purpose was a simple one, SAS wrote in a press release announcing the contract: It will "fight fraud, waste and abuse in the state's unemployment insurance and food stamps program." (Felton, 2015).

In 2020 there was pending litigation against these three companies which led to a range of results (Lawson, 2019; Ethical Tech Initiative 2022).

MiDAS was implemented in October 2013 (AI Now, 2019a: p.1) and the automation of adjudications stopped in late September 2015 (Wykstra, 2020; Shaefer interview, 2020). The UIA was not formally required to cease automating adjudications until January 2017 as part of a state court settlement agreement (Jennifer Lord interview, 2020; Wykstra 2020). However, the seizing of tax refunds and garnishment of wages is still continuing (Lord interview, 2020) and in March 2020, the current director of the University of Michigan Law School's Workers' Rights Clinic testified (Oversight Committee, 2020) to the state Senate Oversight Committee that they believe close to 20,000 of those charged by MiDAS are still being actively pursued and having their wages garnished (Wykstra, 2020).

MiDAS was intended to replace an older system that was 30 years old and to consolidate data and functions that were previously spread over several platforms (Wykstra, 2020).

The Michigan Integrated Data Automated System (MiDAS) determined the eligibility of unemployment insurance (UI) claims by analysing data from past and present claimants, scanning for wage-record irregularities and reporting discrepancies between claimants and their former employers related to the reason for separation from employment (Shaefer and Gray, 2015: p.3). This process was completely automated and when an inconsistency in the record data was detected by the system, that claimant was flagged for potential unemployment fraud and this triggered an email notice to the claimant's UI email address, not their personal email address (Shaefer interview, 2020).

This notice threatened a "determination based on available information," if they failed to respond within ten days (Shaefer and Gray, 2015: p.3). If no response was received the system automatically issued a charge for repayment in addition to a quadruple penalty, the highest fraud penalty in the US at that time (Lord interview, 2020; Shaefer and Gray, 2015: p.6). No other state had a penalty above 150% (AI Now, 2019a: p.2). Further, once a claim was substantiated after a 30-day appeal period, the state could immediately go after a person's wages and federal and state income tax refunds and make a criminal referral if payments weren't forthcoming (Wykstra, 2020). The agency later acknowledged that in the majority of the cases between 2013 and 2015, the system ran from start to finish without any human review (Wykstra 2020).

The quadruple penalty meant that there was tremendous incentive to find fraud and generate funds (Lord, 2020).

One of the key problems with this system as raised by Shaefer, is that the way the MiDAS system detected inconsistencies was flawed and did not match up with the reality of a claimant's work experiences, which were often of an insecure, contractual or temporary nature (see for example Felton, 2015). For example, many people in Michigan claiming UI might work 20 hours one week then zero hours the next week and would report all hours worked on a weekly basis, but the data came in to MiDAS on a quarterly basis and the system was programmed to average the hours across that calendar quarter (Schaefer 2020). This meant the system would average how many dollars were earned across hours worked over three months and use that figure as the weekly average, which was often different to the amounts reported each week by claimants. if the hours a claimant was reporting were inconsistent with that average MiDAS would flag them for fraud. (Shaefer interview, 2020)

40,000 people Michiganers were wrongly accused of fraud between 2013 and 2015 (de la Garza 2020). This was unprecedented and marked a significant increase compared with previous years.

Impact

The Agency reported positive impacts related to the MiDAS system overall, most related to quicker response times and reducing staff time.

MiDAS generated a huge amount of money for the Agency. Money collected from penalties went into a Penalty and Interest (P&I) account. The balance within the P&I account was $3.1 million at the close of FY 2010-11 and grew to $68.8 million as of September 30, 2014. Further, legislation passed in 2015 allowed the state to use the funds – previously only used to support UIA activities and pay for representation for those who couldn't afford it – for other purposes (AI Now 2019, p.1).

An impact assessment was not done before or after the system was implemented. System administrators described it as being developed and deployed "in record breaking time." We could find no mention of trialling or testing the system before implementation.

Legal challenges and reporting of the system have uncovered a wide range of negative impacts by those

who were wrongly identified as fraudulent, including personal trauma, evictions, house foreclosures, homelessness, bankruptcies and ruined credit (Charette, 2018). In some cases getting the Unemployment Insurance Agency to admit fraud allegations were made in error took years and legal challenges, which meant any repayment of wrongly captured funds only barely covered the costs of these errors.

Jennifer Lord, a lawyer representing MiDAS claimants, notes: "These people were falsely accused of fraud. It went on their credit report, it impacted their ability to get loans, to get mortgages for homes, car loans, and when they were approved they were approved at higher rates. People were failing background checks to get jobs. [...] Our argument is yes pay people back the money that you stole from them but on top of that there's got to be some sort of recognition that this decision really injured people and it was completely caused by the State's actions." (Lord interview, 2020).

It would be interesting to understand how Covid has impacted the way MiDAS is used. From brief correspondence with Steve Gray it is our understanding that the Agency has been inundated with new UI claims resulting from Covid, and from interview data it is our understanding that even now the Agency is still experiencing problems with the system. (E.g. see NYT article by Emily Badger in references list).

Going forward

The state continues to use MiDAS and there continues to be a lot of critique. Some civil society organisations argue that the system needs to be scrapped. Legal action continues as those who were wrongly accused and faced significant harms battle for remuneration.

Interviews conducted online in 2020:

Jennifer Lord: Lawyer for Pit McGhee Palmers and Rivers - represented the MiDAS claimants in one of the three lawsuits

Luke Shaefer: Professor of Public Policy and Social Work at the University of Michigan and directs a University wide initiative called Poverty Solutions

Interview requests were also sent to politicians involved with this case study as well as a government

official.

References

AI Now, (2019a) 'Michigan Legislature Factsheet', available: https://drive.google.com/file/d/1GqJ7pHh99Taff5_eP9ziWP_jYDI_sQ_o/view?usp=sharing

AI Now (2019b) 'Litigating Algorithms 2019 US Report: New Challenges to Government Use of Algorithmic Decision Systems' https://ainowinstitute.org/litigatingalgorithms-2019-us.pdf

American Federation of State, County and Municipal Employees (AFSCME) (no date) *AFSCME*. Available at: https://www.afscme.org/ (Accessed: 26 August 2021).

Badger, E. and Parlapiano, A. (2020) 'States Made It Harder to Get Jobless Benefits. Now That's Hard to Undo.', *The New York Times*, 30 April. Available at: https://www.nytimes.com/2020/04/30/upshot/unemployment-state-restrictions-pandemic.html (Accessed: 25 August 2021).

Bailey, L. (2017) *Software Firms Sued Over Mich. Jobless System*, *Courthouse News Service*. Available at: https://www.courthousenews.com/software-firms-sued-michigan-jobless-system/ (Accessed: 26 August 2021).

Charette, R. N. (2018) *Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold*, *IEEE Spectrum*. Available at: https://spectrum.ieee.org/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold (Accessed: 4 August 2021).

'Contract Change Notice' (2018). State of Michigan Enterprise Department.

Cunningham, D. (2015) *Problem Solvers Question Gov. Snyder About Unemployment Insurance Agency Complaints*, *Fox17 West Michigan*. Available at: https://www.fox17online.com/2015/06/18/governor-snyder-questioned-about-unemployment-insurance-agency/ (Accessed: 25 August 2021).

Cwiek, S. (2019) *Democrats Propose Bills to Expand Unemployment Benefits And Help Those Falsely Accused of Fraud*, *NPR Michigan Radio*. Available at: https://www.michiganradio.org/post/democrats-propose-bills-expand-unemployment-benefits-and-help-those-falsely-accused-fraud (Accessed: 25 August 2021).

de la Garza, A. (2020) Automated Systems Trapping Citizens in Bureaucratic Limbo, TIME Magazine. Available at: https://time.com/5840609/algorithm-unemployment/ (Accessed: 20 August 2021).

Demas, S. J. (2017) 'Michigan's Unemployment System Fiasco Has Damaged Too Many Lives', *Susan J. Demas*, 7 February. Available at: http://www.susanjdemas.com/news/2017/2/7/susan-j-demas-michigans-unemployment-system-fiasco-has-damaged-too-many-lives (Accessed: 25 August 2021).

Egan, P. (2019) State of Michigan's mistake led to man filing bankruptcy, Detroit Free Press, 22 December, https://www.freep.com/story/news/local/michigan/2019/12/22/government-artificial-intelligence-midas-computer-fraud-fiasco/4407901002/

Ethical Tech Initiative of DC. The George Washington University. 2022. https://blogs.gwu.edu/law-eti/ai-litigation-database/

Felton, R. (2015) *Criminalizing the Unemployed*, *Detroit Metro Times*. Available at: https://www.metrotimes.com/detroit/criminalizing-the-unemployed/Content?oid=2353533 (Accessed: 20 August 2021).

Fleming, L. N. and Fournier, H. (2015) *Class-Action Lawsuit Filed Over Seized Tax Refunds*, *The Detroit News*. Available at: https://www.detroitnews.com/story/news/local/michigan/2015/09/14/class-action-lawsuit-filed-tax-refunds/72255478/ (Accessed: 26 August 2021).

Lawson, J. D. M. (2019) *Cahoo v. SAS Institute Inc.* Available at: https://www.leagle.com/decision/infdco20190423e59 (Accessed: 20 August 2021).

Levin, S. (2016) 'MiDAS Politician Letter'. Available at: https://drive.google.com/file/d/1CtCt8V5KP-JQGZMCsavhdsomxBCPAHXQ/view?usp=sharing&usp=embed_facebook (Accessed: 20 August 2021).

McFarlane, James on behalf of Michigan Department of Licensing and Regulatory Affairs and Michigan Unemployment Insurance Agency (2014) "Michigan Integrated Data Automated System & Unemployment Insurance Modernization Project" NASCIO award nomination," available: https://web.archive.org/web/20150401113325/http:/www.nascio.org/awards/nominations2014/2014/2014MI7-NASCIO%20Nomination%20-%20Michigan%20Integrated%20Data%20Automated%20System%20(MiDAS)%20&%20Unemployment%20Insurance%20Modernization%20Project3.pdf

MI House Democrats (2019) 'Unemployment Insurance Agency Reform'. housedems.com.

Michigan Department of Licensing and Regulatory Affairs (2013) *Michigan Integrated Data Automated System & Unemployment Insurance Modernization Project*. Unemployment Insurance Agency, Michigan Department of Licensing and Regulatory Affairs, p. 7.

Michigan Eastern District Court (2017) *Zynda et al. v. Zimmer et al.*

News 10 (2019) *House Democrats Have a Plan to Reform Unemployment*, *https://www.wilx.com*. Available at: https://www.wilx.com/content/news/House-Democrats-have-a-plan-to-reform-unemployment-559266021.html (Accessed: 26 August 2021).

Oversight Committee (2020). Lansing: Binsfield Office Building. Available at: https://misenate.viebit.com/player.php?hash=dOK5LThUPBed (Accessed: 20 August 2021).

Pitt, McGehee, Palmer & Rivers (2020) *UIA Fraud Class Action*, *UIA Fraud Class Action*. Available at: http://www.uiafraudclassaction.com/ (Accessed: 26 August 2021).

Richardson, R., Schultz, J. M. and Southerland, V. M. (2019) *Litigating Algorithms 2019 US Report: New Challenges to Government Use of Algorithmic Decision Systems*. AI Now Institute. Available at: https://ainowinstitute.org/litigatingalgorithms-2019-us.html.

Ringler, Doug A. (2016) Michigan Integrated Data Automated System (MiDAS): Unemployment Insurance Agency, Department of Talent and Economic Development and Department of Technology, Management, and Budget. State of Michigan Auditor General. Available at: https://www.audgen.michigan.gov/finalpdfs/15_16/r641059315.pdf

Roelofs, T. (2017) *Broken: The Human Toll of Michigan's Unemployment Fraud Saga*, *Bridge Michigan*. Available at: https://www.bridgemi.com/michigan-government/broken-human-toll-michigans-unemployment-fraud-saga (Accessed: 26 August 2021).

Shaefer, Luke and Gray, Steve (May 2015) 'Michigan Unemployment Insurance Agency: Unjust Fraud and Multiple-Determinations' [Memo to US Department of Labor]

WILX News 10, "House Democrats have a plan to reform unemployment." *WILX News 10* [Online]. September 3, 2019.

Wykstra, S. (2020) *Government's Use of Algorithm Serves Up False Fraud Charges*, *Undark Magazine*. Available at: https://undark.org/2020/06/01/michigan-unemployment-fraud-algorithm/ (Accessed: 26 August 2021).

Zynda, et al. v. Arwood, et al. (2016), https://www.govinfo.gov/app/details/USCOURTS-mied-2_15-cv-11449/summary

Notable legal action

Zynda et al v Zimmer et al (UIA reps). Filed in April 2015, settled Jan 2017. This was the case that resulted (Named Steve Arwood, director of the Department of Talent and Economic Development, and Moffett-Massey as defendants)

Bauserman et al v UIA. Filed in September 2015 (represented by Jennifer Lord) - ongoing. This case is seeking compensation for claimants.

Cahoo et al v SAS Institute, FAST Enterprises LLC, CSG Government Solutions,
Stephen Geskey, Shemin Blundell, Doris Mitchell, Debra Singleton, Julie A. McMurtry, and Sharon Moffet-Massey, Defendants. Filed April 2019 - ongoing.

Robo-debt / Online Compliance Intervention Paused (Australia)

Summary

The Department of Human Services and Centrelink, which manages social security payments in Australia, changed the algorithmic system it used to identify people who had been overpaid benefits in 2016. A new accounting method meant that fortnightly earnings could be used to estimate annual earnings. Also, where previously a human would be responsible for checking a case that was flagged through data matching, the new system was automated which meant debt notices were sent automatically to those who were identified by the system as being overpaid benefits. Thousands more letters were being sent out than had previously been the case. The system went from sending out 20,000 debt notices a year to, at one point, 20,000 a week. Further, it became the responsibility of those receiving debt notices to challenge the notification and prove that the government's system had made a mistake.

After receiving a notice, individuals had 21 days to correct discrepancies logged by the system. Many people sent letters didn't receive them because the letters had been sent to an old address or to myGov accounts no longer checked. This resulted in many people learning they had been flagged by the system when they were contacted by a debt collector (Pett and Briefing, 2017).

Also, challenging a debt notice was difficult as it involved trying to get an officer on the phone which could take hours and finding paperwork, sometimes going back 7 years. Those receiving the letters reported their frustration on a website set up by a coalition to challenge the new system called NotMyDebt. One social service organisation reported that a quarter of the debt notices it investigated were wrong. The automated debt recovery system was labelled RoboDebt by critics.

Why was it stopped?

The system faced considerable pressure from civil society organisations, activists, and politicians. Victoria Legal Aid challenged the system in court and in November 2019 a judge ruled that the system was unlawful (Medhora, 2019). The government paused (Barbaschow, 2019) the automated debt-matching component in November. In May 2020 the Government announced it would repay all debts at a cost of $721 million. A Senate Committee has recommended the system be terminated. A class action lawsuit was settled in November 2020 for $1.2 billion. It is unclear if the system will be started again in a different form.

Key factors leading to change

The system did not work, thousands of debt notices were sent out in error to those with little resources to challenge the government. This led to widespread mobilization which included those affected, community organisations and legal aid and legal challenges. A campaign called Not My Debt was initiated which contributed to critical media coverage and effort was made to mobilize politicians. This mobilization included, but was not limited to, activists such as Asher Wolf, the Australian Council of Social Service (ACOSS), the Australian Unemployed Workers Union, Economic Justice Australia, and Victoria Legal Aid. The Green Party and the Labour Party also were critical.

This effort led to political and bureaucratic review and critique including a Commonwealth Ombudsman Report (2017), a report by the Australian National Audit Office (2017), and two Senate Inquiries (2017 and 2020). Ultimately, a legal challenge brought by Victoria Legal Aid (Victoria Legal Aid, 2017) resulted in a court decision that the system was unlawful and a class action lawsuit brought by Gordon Legal led to

a settlement in November 2021 (Gordon Legal, 2021).

How did the system work?

In Australia the Centrelink Online Compliance Intervention system is responsible for a number of payments from the government. In 2016 the part of the system used to identify who had been overpaid benefits was changed. Previously a data matching system had been used to compare reported earnings with tax office data and when a mismatch was found a human was responsible for looking into the case to identify if someone had been overpaid benefits.

There were two key changes made in 2016. The first was a change to the accounting method used to calculate earnings which meant that fortnightly income could be averaged to estimate a person's income for an entire year. This was a problem for people whose income fluctuates, for example people on contracts, people who work part-time or who have precarious or seasonal work. The second key change was that a human was no longer responsible for overseeing an investigation to determine if there had been an overpayment. The new system was automated so that a debt-notice was sent out when a case was flagged. The calculations could go back five years. The system went from sending out 20,000 debt notices a year to at one point sending out 20,000 a week.

Once receiving a debt-notice it was the responsibility of those receiving them to prove the notice was incorrect which in some cases would mean tracking down old pay stubs or trying to find out from Centrelink how the over-payment had been calculated.

A coalition was formed made up of those affected (who had a range of skill sets), activists, community organisations and lawyers. They created the NotMyDebt (#NotMyDebt) webpage to share their experiences. Some people reported spending hours trying to speak with someone at Centrelink to find out why they had been sent this debt notice, others reported being unable to find out why the decision had been reached even after talking to someone at Centrelink.

There were thousands of people sent notices that were wrong. The federal government has admitted that almost 400,000 people have paid 470,000 debts worth $721m that were not owed (Phillips, 2020).

This came after a legal challenge. In November 2019 the federal court settled a challenge to its Robo-debt program, conceding that a $2,500 debt raised against Deanna Amato was not lawful because it relied on income averaging (Karp 2019). Around this time *sole reliance* on income averaging was stopped. In November 2019, Services Australia (then named Department of Human Services, and which Centrelink is a part of) wrote: "The department has made the decision to require additional proof when using income averaging to identify over payments. This means the department will no longer raise a debt where the only information we are relying on is our own averaging of Australia Taxation Office income data" (Farrell 2019).

Impact

The Department of Human Services (now Services Australia), of which Centrelink is a part, conducted a pilot to inform the design of the online compliance system.

No impact assessment was done before implementing the system and this was a critique voiced in the Senate inquiry. The Commonwealth Ombudsman published a report in 2017 that was critical of the system and noted the need for there to have been more work to explain the debt averaging and in a later report was also critical of the government for not checking the legality of the system.

Impact was discussed in the first Senate Inquiry where those affected by the errors reported how they were negatively affected. People reported the stress, pressure and trauma caused by the notices and debt collection (done in error) (Senate 2020, p. 16).. In some cases this was done to people with mental health issues, to people with poor health conditions and struggling with illness and disease and to those living in precarious situations. People reported the negative psychological affects of these notices and also of the strain of having a payment withheld (Henriques-Gomes, 2020; Senate 2020, p. 16).

One of the organisers of the Not My Debt campaign has described Robodebt as an "algorithmic weapon of calculated political cruelty" because of the negative impact it had on those falsely accused and the way it disproportionately targeted "the unemployed, disabled people, single parents, care-givers, casual and gig economy workers" (Wolf, 2020).

Going forward

There have been calls for a Royal Commission to investigate this case fully. The settlement for the class action lawsuit is expected. Robodebt has been described as "fundamentally a public policy failure". An example of a failure which used technology, but "the failure was upstream of the technology, the failure was in the inception of the system itself" (Townsend interview, 2020).

Concerns have been raised that the refunds people are entitled to are not being returned very quickly (Henriques-Gomes 2020).

A system called Single Touch Payroll is being rolled out and it has been developed with the goal of keeping better track of what people earn.

Sources

Media: ABC (Pett and Briefing, 2017), Guardian (Henriques-Gomes, 2019), Guardian (Karp, 2019), Guardian (Karp, 2019), Guardian (Henriques-Gomes, 2020)

Research: Carney, 2018

Civil Society: #NotMyDebt

Government: Ombudsman Review, First Senate inquiry, Second Senate Inquiry

Interviews conducted online, 2020:

Joel Townsend (Program Manager, Economic and Social Rights, Victoria Legal Aid)

Rachel Siewert (Senator for Western Australia 2005-2021, Australian Greens)

References

Australian Government (2021) Information about refunds for the Income Compliance Program, Service Australia. Available at: https://www.servicesaustralia.gov.au/individuals/subjects/information-about-refunds-income-compliance-program#whichdebtrepayments

Barbaschow, A. (2019) *Government Backflip as Robo-Debt Income Automation Paused*, *ZDNet*. Available at: https://www.zdnet.com/article/government-backflip-as-robo-debt-income-automation-paused/ (Accessed: 20 August 2021).

Carney, T. (2019) 'Robo-Debt Illegality: The Seven Veils of Failed Guarantees of the Rule of Law?', *Alternative Law Journal*, 44(1), pp.

4–10. doi: 10.1177/1037969X18815913.

Community Affairs References Committee (2017) Design, scope, cost-benefit analysis, contracts awarded and implemented with the Better Management of the Social Welfare System initiative, The Senate, June. Available at: https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/SocialWelfareSystem/Report

Commonwealth Ombudsman (2019) *Submission by the Commonwealth Ombudsman: Inquiry Into Centrelink's Compliance Program*.

Farrell, P. (2019) Government halting key part of robodebt scheme, will freeze debts for some welfare recipients. ABC News. Available at: https://www.abc.net.au/news/2019-11-19/robodebt-scheme-human-services-department-halts-existing-debts/11717188

Glenn, R. (2017) Centrelink's automated debt raising and recovery system, Commonwealth Ombudsman, Report No. 02/2017. Available at: https://www.ombudsman.gov.au/__data/assets/pdf_file/0022/43528/Report-Centrelinks-automated-debt-raising-and-recovery-system-April-2017.pdf

Gordon Legal (2021) *Robodebt Class Action Settlement*, *Gordon Legal*. Available at: https://gordonlegal.com.au/robodebt-class-action/ (Accessed: 20 August 2021).

Henriques-Gomes, L. (2020) All Centrelink debts raised using income averaging unlawful, Christian Porter concedes, *Guardian*, available at: https://www.theguardian.com/australia-news/2020/may/31/all-centrelink-debts-raised-using-income-averaging-unlawful-christian-porter-concedes

Henriques-Gomes, L. (2019) *Centrelink Still Issuing Incorrect Robodebts to Meet Targets, Staff Claim*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2019/may/29/centrelink-still-issuing-incorrect-robodebts-to-meet-targets-staff-claim (Accessed: 20 August 2021).

Henriques-Gomes, L. (2020a) *Officials Refuse To Say How Many Flawed Robodebts They Have Found*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2020/mar/06/centrelink-robodebt-senate-estimates-class-action (Accessed: 12 August 2021).

Henriques-Gomes, L. (2020b) 'Robodebt-related trauma': the victims still paying for Australia's unlawful welfare crackdown,' The Guardian, available: https://www.theguardian.com/australia-news/2020/nov/21/robodebt-related-trauma-the-victims-still-paying-for-australias-unlawful-welfare-crackdown

Karp, P. (2019a) *Government Admits Robodebt Was Unlawful As It Settles Legal Challenge*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2019/nov/27/government-admits-robodebt-was-unlawful-as-it-settles-legal-challenge (Accessed: 12 August 2021).

Karp, P. (2019b) *Robodebt: Government Abandons Key Part Of Debt Recovery Scheme In Major Overhaul*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2019/nov/19/robodebt-government-abandons-key-part-of-debt-recovery-scheme-in-major-overhaul (Accessed: 12 August 2021).

Medhora, S. (2019) *Federal Court Rules Robodebt 'Unlawful' in Significant Court Case*, *triple j*. Australian Broadcasting Corporation. Available at: https://www.abc.net.au/triplej/programs/hack/federal-government-loses-major-robodebt-case/11742494 (Accessed: 20 August 2021).

#NotMyDebt (no date) *#NotMyDebt*, *#NotMyDebt*. Not My Debt. Available at: https://www.notmydebt.com.au/ (Accessed: 20 August 2021).

Parliament of Australia, C. (no date a) *Centrelink's Compliance Program*. Available at: https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/Centrelinkcompliance (Accessed: 26 August 2021).

Parliament of Australia, C. (no date b) *Design, Scope, Cost-Benefit Analysis, Contracts Awarded and Implementation Associated with the Better Management of the Social Welfare System Initiative*, *Parliament of Australia*. Available at: https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/SocialWelfareSystem (Accessed: 26 August 2021).

Pett, H. and Briefing, C. C. for B. (2017) 'We're All Talking About Centrelink, but What is "Robodebt" Anyway?', *ABC News*, 2 March. Available at: https://www.abc.net.au/news/2017-03-03/centrelink-debt-controversy-what-is-robodebt/8317764 (Accessed: 20 August 2021).

Phillips, J. (2020) Robodebt was a flagrant abuse of government power. It should never have seen the light of day." *Guardian*, 29 May. Available at: https://www.theguardian.com/commentisfree/2020/may/29/robodebt-was-a-flagrant-abuse-of-government-power-it-should-never-have-seen-the-light-of-day

Senate Standing Committee on Community Affairs (2020) Centrelink's compliance program, September, https://parlinfo.aph.gov.au/parlInfo/download/committees/reportsen/024338/toc_pdf/Centrelink'scomplianceprogram.pdf;fileType=application%2Fpdf.

Victoria Legal Aid (2017) *Robo-Debts*. Victoria Legal Aid. Available at: https://www.legalaid.vic.gov.au/find-legal-answers/centrelink/robo-debts (Accessed: 20 August 2021).

Victoria Legal Aid (2019) Notice of Filing, VID611/2019. Available at: https://www.legalaid.vic.gov.au/sites/www.legalaid.vic.gov.au/files/statement_of_agreed_facts.pdf

Wolf, A. (2020) Robodebt was an algorithmic weapon of calculated political cruelty. The Canberra Times. Available at: https://www.canberratimes.com.au/story/6775350/robodebt-was-an-algorithmic-weapon-of-calculated-political-cruelty/

UK Local Authorities Stop Using Risk Based Verification System

Summary

Several UK Local Authorities have stopped using Risk Based Verification systems to process benefits claims. The systems are often external systems used to classify the level of fraudulent risk of claimants as low, medium and high. Low risk cases require little verification while more checking is required when someone is ranked high risk. The categorizations of risk "are confidential to the supplier so local authorities do not know why cases are categorised as they are."

Why was it stopped?

Local authorities have reported that the system is not delivering anticipated cost savings or improving processing times. Some councils have raised concerns about accuracy and the lack of transparency of company systems. It has also been reported that the introduction of a new system, Universal Credit, is leading some local authorities to stop using RBV.

Bristol City Council began using RBV for new claims and changes in October 2014. They state in September 2020, the point at which they halted their use of the system:

It has not delivered the anticipated savings in workload for staff or significant improvements in average processing times. (Bristol City Council, 2020, page 1)

Additionally, they state that despite RBV's promoted advantage of redirecting resources towards "high risk" cases, it identified "very few fraudulent cases and the management information available has offered no assurance of the value of this process." They also cite significant changes in the wider welfare landscape and new technologies. Such changes include the roll out of Universal Credit (which has significantly reduced the number of new claims for Housing Benefit) and new systems such as the Verify Earnings and Pensions (VEPS) system "to check current earnings and pension data provided to HMRC by employers/ pension providers", and the DWP's Housing Benefit Matching Service and access to its Searchlight system "which allows validation of almost all DWP/HMRC benefits" (Bristol City Council, 2020, page 1).

The Guardian reported that North Tyneside Council had stopped using their TransUnion RBV system due to delays incurred by the system when it incorrectly identified low risk claims as high risk (Marsh, 2019). A document report by the council highlighted that the system "provides no reason for a case meeting a high-risk category and it was found that in most cases the reason for it being high risk could not be established" (North Tyneside Council, 2019, page13). Harrow Council withdrew the use of RBV for Housing Benefit and Council Tax Support assessments in 2020 citing the rollout of Universal Credit "the level of complexity of a high proportion of residual claims" making RBV less effective (Harrow Council, 2020).

Key factors leading to change

Discussion within local authorities about the limits of the system have been raised as a key factor leading to change. In addition it was suggested that the introduction of the new Universal Credit system also played in role in leading local authorities to stop using RBV.

How does it work?

Risk Based Verification (RBV), as it is used in UK benefits contexts by councils, is a system for risk scoring

claimants' applications to justify increased or decreased scrutiny. The Department of Work and Pensions recommended that local authorities start using RBV in 2011 and outlined what could qualify as low, medium and high-risk claims (DWP, 2011, paragraph 9). These included:

Low Risk Claims: Only essential checks are made, such as proof of identity. Consequently these claims are processed much faster than before and with significantly reduced effort from Benefit Officers without increasing the risk of fraud or error.

Medium Risk Claims: These are verified in the same way as all claims currently, with evidence of original documents required. As now, current arrangements may differ from LA [Local Authority] to LA and it is up to LAs to ensure that they are minimising the risk to fraud and error through the approach taken.

High Risk Claims: Enhanced stringency is applied to verification. Individual LAs apply a variety of checking methods depending on local circumstances. This could include Credit Reference Agency checks, visits, increased documentation requirements etc. Resource that has been freed up from the streamlined approach to low risk claims can be focused on these high risk claims.

The goal of using RBV seems to be to speed up application processing times and to increase fraud detection. Council notes indicate that these systems make use of 40 or 50 variables in their predictions.

The systems are proprietary and it is unclear how the risk calculations are done.

Impact

Concerns have been raised about the transparency of the systems, as many local authorities rely on private companies for these systems. The UN Special Rapporteur on extreme poverty and human rights raised concerns about a lack of transparency and about the accuracy of the predictions. "The presumption of innocence is turned on its head when everyone applying for a benefit is screened for potential wrongdoing" (Alston, 2018).

MedConfidential raises concerns about process and the ability of people to challenge scores they are not aware of as well as how someone can move from high risk to a lower category. Concerns have also been raised about the ability of local authorities to assess the services being provided by private companies when little information is provided about how the system works. Researchers have raised concerns about an inability for public evaluation of the impact and fairness of these systems (Harris, 2020).

Going forward

It has been reported that the DWP is "developing a "fully automated risk analysis and intelligence system for fraud and error," which will go beyond automatically finding inconsistencies between different databases and aims to prevent fraud and error by using new tools including Artificial Intelligence" (Alston. 2018).

Sources

Interviews: Two local authority employees (anonymous)

References

Alston, P. (2018) *Statement on Visit to the United Kingdom, by Professor Philip Alston, United Nations Special Rapporteur on extreme poverty and human rights*, *United Nations Human Rights Office of the High Commissioner*. Available at: https://www.ohchr.org/en/NewsEvents/Pages/DisplayNews.aspx?NewsID=23881 (Accessed: 18 August 2021).

Berry, K. & Shaw, J. (2017) "Council Tax Reduction". Scottish Parliament, Research Briefings and Fact Sheet.

Bristol City Council (2018) "Risk Based Verification Policy", revised 27th March 2018. https://democracy.bristol.gov.uk/documents/s51651/Appenddix%20A.3%20RBV%20Policy.pdf Archived copy: https://web.archive.org/web/20201211161402/https://democracy.bristol.gov.uk/documents/s51651/Appenddix%20A.3%20RBV%20Policy.pdf

Bristol City Council (2020) "Review and rationale for replacement of RBV", 1st September 2020. https://democracy.bristol.gov.uk/documents/s51649/Appendix%20A.1%20Review%20and%20rationale%20for%20replacement%20of%20RBV.pdf Appendix A.1 of the proceedings of Tuesday 1st September 2020's Cabinet session. https://democracy.bristol.gov.uk/ieListDocuments.aspx?CId=135&MId=8399 Further decision details: https://democracy.bristol.gov.uk/ieDecisionDetails.aspx?Id=862

Chichester District Council (2017) "Risk Based Verification Policy 2017". https://chichester.moderngov.co.uk/documents/s10306/Housing%20Benefit%20and%20Council%20Tax%20Reduction%20Risk%20Based%20Verification%20Policy%202018-2019%20-%20Appendix.pdf Archived copy: https://web.archive.org/web/20200305123427/https://chichester.moderngov.co.uk/documents/s10306/Housing%20Benefit%20and%20Council%20Tax%20Reduction%20Risk%20Based%20Verification%20Policy%202018-2019%20-%20Appendix.pdf

Harris, Swee Leng (2020) "Data Protection Impact Assessments as rule of law governance mechanisms", *Data & Policy*, 2: e2: 1-12. https://www.cambridge.org/core/services/aop-cambridge-core/content/view/3968B2FBFE796AA4DB0F886D0DBC165D/S2632324920000036a.pdf/data_protection_impact_assessments_as_rule_of_law_governance_mechanisms.pdf

Marsh, S. (2019) "One in three councils using algorithms to make welfare decisions", *The Guardian*, 15th October 2019. https://www.theguardian.com/society/2019/oct/15/councils-using-algorithms-make-welfare-decisions-benefits

McKeever, G. (2017) "Social security devolution: Northern Ireland and Scotland" [briefing paper]. http://www.parliament.scot/S5_Social_Security/Inquiries/112._McKeever_Professor_Grainne.pdf

MedConfidential (2020) "Annex 2: Risk Based Verification", September 2020. https://medconfidential.org/wp-content/uploads/2020/09/Annex-2-RBV-20200908.pdf

Policing ADS

Overview

The policing case studies present different approaches to the use of ADS, particularly as related to whether or not there is pre-implementation testing, considerations of impact, response to community concerns and effort to seek independent review. Despite the differences in approaches, thematically these case studies bring to the fore questions about care and effectiveness.

Policing

Germany, Baden-Württemberg	The German federal state of Baden-Württemberg stops using PRECOBS predictive policing system (2015-2019)
United States, Los Angeles	Los Angeles Police Department stops using Los Angeles Strategic Extraction and Restoration (Laser) (2011-2018) and PredPol (2009-2020)
New Zealand	The New Zealand High Tech Crime Group decides not to pursue use of Clearview facial recognition technology.
UK, Durham	Durham police stop using the Harm Risk Assessment Tool (HART) (2016-2020)

These case studies demonstrate the differing ways that *care* can be identified in different contexts, as in one case care presents through the use of careful piloting and review and in another case as community mobilization in response to concerns about negative impact. In the case of the German federal state of Baden-Württemberg's trial of PRECOBS care presents, in practice, as scepticism, time taken to do careful investigation, the engagement of a meaningful independent review process and the transparency of this process. In the case of the Los Angeles Police Department two ADS were cancelled, LASER and PredPol. Care in this context can be identified in the decade-long mobilizing efforts of the Stop LAPD Spying Coalition to raise concerns and questions about the impact and effectiveness of these systems as well as the work done to document impact and raise concerns about how these systems exacerbate systemic injustice and discrimination.

Delivering on intended outcomes

The policing case studies, like those in the area of fraud detection, bring to the fore questions about *effectiveness*. These case studies challenge the rationale that ADS will always lead to more efficient services and a reduction in costs. The police trial of PRECOBS in Baden-Württemberg led to the conclusion that the system is useful for statistics but not very good at predicting where burglary would occur. The LAPD systems cancelled demonstrate the importance of ensuring efforts to measure effectiveness and impact. In this case the Office of the Inspector General noted the difficulty tracking the effectiveness of the system given inconsistencies in data collection and implementation. In New Zealand, the police force found the use of facial recognition technology not effective and decided not to pursue its use after trying the technology, although recent news coverage suggests it may now be pursuing new tools. The issue of accuracy also presents in reviews of the use of the Harm Risk Assessment Tool by the Durham Police in the UK.

Germany – PRECOBS Case Study

Summary

The German federal state Baden-Württemberg trialled PRECOBS, a predictive policing system for 4 years in the cities Stuttgart and Karlsruhe (2015-2019). The system was trialled but never implemented. This was decided after an evaluative study by the Max-Planck Institute for the Study of Crime, Security and Law that found disappointing results in terms of reducing the number of burglaries through PRECOBS.

Why was it not implemented?

The police forces in Stuttgart and Karlsruhe reported that the system was found as not efficient enough after trialling it for four years. They argue that PRECOBS is useful for statistics, but not very good at making predictions because it only captures specific types of professional burglary, while many others are not considered. Moreover, an evaluative study by the Max-Planck Institute concluded that PRECOBS's usefulness is highly dependent on the area it is applied in, and that for many areas, it is not very useful for predictions.

Key factors leading to change

The government and the police force involved were clear from the beginning that they were piloting the use of this system in order to research its use. They also sought an independent external review. The cautious research approach in this case as well as research findings led to the government decision not to pursue implementation. It was decided that resources could be more effective if directed to other areas of policing.

How did the system work?

The goal was to see if this near repeat prediction software is able to predict and prevent domestic burglaries and if it can be tailored and embedded into daily police work. It was noted that from the outset this was intended as a scientific study, to see if the system would be useful to police

PRECOBS is used to predict the likelihood of future burglaries and is based on the observation that crime events are followed by further events in spatial and temporal proximity. It is based on the idea that burglars operate rationally and will try to obtain the maximum resources with minimum effort and risk. The police worked with the Institut für musterbasierte Prognosetechnik (IfmPt) to trial PRECOBS.

The system identifies events which are likely to be followed by near-repeat burglaries. The idea was that information about where near events are likely to take place could be used in operational planning. The software uses data from the past (usually the last five years) in order to identify which attributes of residential burglaries indicate there will be near-repeats, where a follow-up offence is likely. The system relies heavily on the circumstances of an offence and the geographic location, what is called triggers. The system also identifies anti-triggers. Details like stolen goods, method of entry, tools and locality are considered as are areas with a high proportion of near repeat burglaries in the past. Concerned about privacy protection, the police force consulted with their data protection officer who said that housing data is personal data so they took steps to anonymize housing data.

The entire trial is estimated to have cost 500,000 €.

During the pilot, police data about burglaries were imported to PRECOBS three times a day. Attributes of the recent burglary events were compared to the trigger catalogues. If an attribute matched the trigger criteria and the burglary happened in an area with near-repeats, then an automated prediction was made. The operator would check this and if accepted an alert was sent to the local police with a PDF document. The message had a map, additional information and best times to patrol. The area is identified as at a heightened risk for near repeat burglaries for 7 days. The system only works with human oversight.

It has been stressed that this predictive system was not a black box and not complex, that the algorithm was easy to understand.

The State Office of Criminal Investigations stressed that there would be an external evaluation of the project. This evaluation was done by the Max-Planck Institute for the Study of Crime, Security and Law. Some viewed this as an impact assessment. It was stressed that the police in Baden-Württemberg sought for this work to be evidence based and that they had a difficult time finding independent evaluations of predictive policing systems.

Impact

There was criticism from police officers that a system led to more work as alarms meant that more presence was required in some areas than others. In general, the external review found that half of the officers supported the system and half were opposed to it. Very few were neutral.

There were two pilot phases to this project with reviews of both pilots. These reviews suggest that the system is moderately effective in urban areas, but not effective in more rural areas. Operators and analysts said that the dashboard made it easier for them to get an overview of the data and to analyse it. Operators also reported that the system made it easier to convince colleagues about why they should go to a particular neighbourhood, their colleagues were more likely to accept their suggestion if backed by the system. Some negative effects reported include a decrease in a local police station's autonomy, increased workload, it 'binds forces' that may be better used elsewhere and the cost of the software.

Given the high costs, the system was deemed by the external review and the police force to not be efficient enough for the Baden-Württemberg police. They concluded that the system can have some impact, but that resources could be better used by investing in other areas of police work. The State Interior Minister Thomas Strobl followed this recommendation and stopped the system use and the research project.

There was little criticism from civil society organisations and researchers. It is thought this is due, in part, to the decision to consult with the data protection officer from the outset, to seek an external independent review and to make much information publicly available. Media coverage has been mostly positive.

Going forward

There is interest in using software to better understand the overview of a situation and the connections between location, time and the nature of crime. It has been said work is being done in relation to this, but not prediction. Instead, they are focusing on the root causes of crime in analysis.

Interviews conducted online, 2020:

Dominik Gerstner, Research fellow at the Max-Planck Institute for the Study of Crime, Security and Law (MPI)

Interview A, Government

References

Datenschutz (2014-2015) Tätigkeitsbericht des Landesbeauftragten für den Datenschutz Baden-Württemberg 2014/2015. Available at: https://www.baden-wuerttemberg.datenschutz.de/wp-content/uploads/2016/02/32._TB.pdf

FragDenStaat (2017) Auskunft über PRECOBS Predictive Policing Software Datensets und Gewichtung der Faktoren. Available at: https://fragdenstaat.de/anfrage/auskunft-uber-precobs-predictive-policing-software-datensets-und-gewichtung-der-faktoren/67030/anhang/20170519-lifg-anfrage-precobs_geschwaerzt.pdf

Gerstner, D. (2017) Predictive Policing als Instrument zur Prävention von Wohnungseinbruchdiebstahl. Available at: https://pure.mpg.de/rest/items/item_2498917_4/component/file_3014304/content

Gerstner, D. (2018) Predictive Policing in the Context of Residential Burglary: An Empirical Illustration on the Basis of a Pilot Project in Baden-Württemberg, Germany, European Journal for Security Research, 3: 115-138. Available at: https://pure.mpg.de/rest/items/item_3002442_6/component/file_3046246/content

Hammel, EM (2017) Entzauberte Vorhersage, heise online. Available at: https://www.heise.de/hintergrund/Entzauberte-Vorhersage-3852041.html

Heuberger, S. (2017) Predictive Policing: Software ist keine Wunderwaffe gegen Einbrüche, GQ. Available at: https://www.gq-magazin.de/auto-technik/article/mit-big-data-den-einbrecher-schnappen-klappt-nur-bedingt

IfmPt (2019) PRECOBS in Baden-Württemberg. Available at: https://www.ifmpt.de/project_baden-wuerttemberg.html, Max-Planck-Institut.

Klingbeil, J.

Mayer, N. (2017) Software soll Einbrecher stoppen, StN.de. Available at: https://www.stuttgarter-nachrichten.de/inhalt.land-testet-einbruchs-software-kommissar-computer.ee56faed-3723-405d-8986-9aa76dfc888d.html

Krasmann, S. and Egbert, S. (2019) Predictive Policing. Eine ethnographische Studie neuer Technologien zur Vorhersage von Straftaten und ihre Folgen für die polizeiliche Praxis, Universität Hamburg. Available at: https://www.wiso.uni-hamburg.de/fachbereich-sowi/professuren/hentschel/forschung/predictive-policing/egbert-krasmann-2019-predictive-policing-projektabschlussbericht.pdf

Mayer, N. (2019) Strobl entscheidet sich gegen Precobs, StN.de. Available at: https://www.stuttgarter-nachrichten.de/inhalt.aus-fuer-die-einbruchvorhersage-software-strobl-entscheidet-sich-gegen-precobs.19a18735-9c8f-4f1a-bf1b-80b6a3ad0142.html

Ministerium des Inneren, für Digitalisierung und Kommunen Baden-Württemberg (2015) Polizei Baden-Württemberg startet Einsatz der Prognose-Software "precobs". Available at: https://im.baden-wuerttemberg.de/de/service/presse-und-oeffentlichkeitsarbeit/pressemitteilung/pid/polizei-baden-wuerttemberg-startet-einsatz-der-prognose-software-precobs/

ntv (2019) Gewerkschaft kann Aus für Precobs-Software nachvollziehen, September. Available at: https://www.n-tv.de/regionales/baden-wuerttemberg/Gewerkschaft-kann-Aus-fuer-Precobs-Software-nachvollziehen-article21250215.html

LA Police Department stops using LASER and PredPol predictive policing programmes

Summary

The Los Angeles Police Department (LAPD) began using Los Angeles Strategic Extraction and Restoration, or Operation LASER, in 2011. It was developed by Palantir. The LAPD was an early adopter of predictive policing and started using PredPol in 2009. PredPol is an algorithm developed by the LAPD in collaboration with local universities.

Why were they stopped?

In 2018 the Stop LA Police Dept Spying Coalition made a demand in its critical report of predictive policing "Before the bullet hits the body" that the Office of the Inspector General (OIG) should conduct a review of data driven policing strategies used by the LAPD. This was heeded by the OIG and in March 2019 the report from Inspector General Mark Smith's internal audit pointed to a number of problems including lack of oversight and inconsistent criteria used to predict crime. The audit also raised concerns about how suspects were racially identified.

According to Muckrock, LASER was suspended in August 2018. But this was not discovered until Spring 2019 as the LAPD did not publicly share this information until then. The LA Times reports (Puente, 2019) that "the move came after a meeting Tuesday [9th April 2019] at which members of the department's civilian oversight panel questioned the effectiveness of data-driven strategies." The LA Times also writes that Josh Rubenstein, the LAPD's chief spokesman, said "We discontinued LASER because we want to reassess the data. It was inconsistent. We're pulling back."

In April 2020, the LAPD announced that they would stop using PredPol. It was reported in BuzzFeed (Haskins, 2020) that Police Chief Michael R. Moore said the police would stop using PredPol due to COVID-19 related financial constraints. Campaign coordinator for the Stop LAPD Spying Coalition said he thinks the group's organizing prompted the LAPD to stop using PredPol.

Key factors leading to change

Community mobilization was a key factor leading to change in the case of LASER and PredPol being stopped.

The Stop LAPD Spying Coalition campaigned (Garcia, 2020) against PredPol and LASER since 2010. They engaged the media, employed academic pressure and filed a lawsuit to access public records, but they argue that community power has been the most important. The Coalition organised protests and direct action within the community, for example by ensuring a presence at weekly Commission meetings. Khan argues that presence at these meetings presented an opportunity to ask key questions about the systems that were not being asked, such as about impact and effectiveness (Khan interview, 2020).

The Coalition tried to access as many documents as possible about PredPol and LASER through public records requests. They did research and community education. They targeted actions to mobilize UCLA staff and students. The ACLU as well as UCLA staff and students challenged these systems.

The Stop LAPD Spying Coalition called for the Office of the Inspector General review of LASER which led

to an internal audit that was critical. A civilian oversight panel was also critical of the effectiveness of the system. The police say COVID 19 and the financial constraints surrounding it was a key factor leading them to stop using PredPol, members of the Coalition argue that years of community protest influenced their decision.

How did these systems work?

LASER was made possible with funding from the Bureau of Justice Assistance, which is part of the Department of Justice. The BJA continually offers funding to different law enforcement agencies and invites them to apply for grants. (Garcia interview, 2020).

Both PredPol and LASER are also a part of the Strategically Managed, Analysis and Research-driven Technology based (SMART) Policing Initiative, which was funded by the Bureau of Justice Assistance (SLSC, 2018: p.6). The LASER grant application states that the project period was expected to take two years to develop with a budget of $500,000 (LAPD, 2009: p.1).

PredPol

PredPol is a private company based in Santa Cruz, California. It was co-founded by Jeff Brantingham, a Professor of Anthropology at UCLA (PredPol 'About' web page, 2020). Initially, PredPol offered its software to LAPD for free in return for LAPD crime data; it was piloted by the LAPD.

PredPol is designed to "predict" where and when crimes will most likely occur over the next 12 hours (Smith, 2019: p.25). The PredPol software uses an algorithm that analyzes 10 years of crime data, including the types of crimes, as well as their locations, dates, and times. PredPol results are generated by the software›s algorithm, and its data does not include information about specific individuals (ibid). According to the OIG, the LAPD used PredPol for two categories of vehicle-related crimes – Motor Vehicle Theft and Burglary/Theft from a Vehicle.

According to SLSC the ten years of reported crime data used by PredPol is from community members' calls for police service and from patrol officers' crime reports. Using the three measures of crime type, time and location, PredPol makes statistically-driven predictions as to which 500 x 500 square foot areas in Los Angeles, called "hot spots" (and sometimes referred to as PredPol boxes) have the highest expected crime rate within the city. The LAPD produces reports at the beginning of each shift highlighting which hotspots have been targeted by PredPol's procedure; these reports are then distributed to officers as a guideline for their patrols (SLSC, 2018: p.7).

As the SLSC explain in their report on predictive policing: "PredPol's model is adapted from a model used to predict clusters of earthquake aftershocks; the creators of the technology allege that 'crime is often generated by structures in the environment, like a high school, mall parking lot or bar'; they compare these institutions to distressed fault lines in their effect on successive events." (ibid: p.8). Also, according to early studies of PredPol (Ferguson, 2016), it focused on predicting property crimes like burglary, in-car theft, and car theft, whereas LASER was more focused on violent crimes and especially gun-related crime (Naguib interview, 2020).

LASER

LASER was developed in house in partnership with Justice and Security Strategies (a consulting firm). There was also some corporate involvement with Palantir, it was the creator of Mission Sheets, one tool used as part of the LASER program.

In the 2009 LASER grant application (Stop LAPD Spying Coalition, 2018) LAPD claim they want to reduce gun violence as part of the Smart Policing Initiative. A Bureau of Justice report on smart policing notes: "The basic premise is to target with laser-like precision the violent repeat offenders and gang members who commit crimes in the specific target areas. The program is analogous to laser surgery, where a trained medical doctor uses modern technology to remove tumors or improve eyesight. First, the area is carefully diagnosed: Who are the offenders, and where and when are they involved in criminal activity? Plans are then developed to remove offenders from an area with minimal invasiveness and minimal harm to the people and areas around them. Extraction of offenders takes place in a 'non-invasive' manner (no task forces or saturation patrol activities), and the result produces less disruption in neighborhoods. Continuing with the medical analogy, by extracting offenders surgically, recovery time of the neighborhood is faster." (Bureau of Justice Assistance, 2012, cited in Smith, 2019: p.4)

LASER had two components that both responded to gun and gang violence: a person-based and a place-based predictive policing strategy. The person-based strategy is referred to as the Chronic Offender Program. The location-based strategy focuses on identifying and increasing police presence in hotspots referred to as LASER Zones (Smith, 2019: p.4).

To implement the LASER program, divisions are required to develop a Crime Intelligence Detail (CID). The CID is composed of three sworn officers and one crime analyst and is responsible for overseeing the LASER program in their respective divisions. Hot spots, also known as LASER Zones, are created in targeted neighborhoods within a division. LASER Zones are different from PredPol's "hot spots," in that they are created by using a mapping system called ArcGIS that analyzes locations of reported crime, arrest data, and calls for service that correlate to gun violence or a violent crime. LASER Zones are maintained for 9-12 months (SLSC 2018: p.9).

Once a division creates a targeted area the CID develops what is called the Chronic Offender Bulletin (COB), which is the program's person-based component (OIG, 2019: p.1). A COB is like a "Most Wanted" poster; however, unlike a "Most Wanted" poster in which a person is formally charged with a crime, a chronic offender is designated a "person of interest." The COB marks individuals for surveillance and a range of potential interventions. (SLSC, 2018: p.10)

Using an amalgamation of information from the COB, Field Interview cards, custody reports, crime reports and arrest reports, a crime analyst then performs an initial screening of community members on the COB and passes them on to the "work up" stage (ibid). This stage utilises Palantir software to access multiple databases "to track and trace any activity related to that person over the previous two years" including a person's criminal history, gang affiliation, previous detentions, and other associations (ibid). Reviewing all this information is called a "work up" of chronic offenders, but analysts also review an individual's physical characteristics, parole or probation status, and locations where the individual has been previously stopped (ibid).

Once "worked up" in Palantir, individuals are assigned points based on five weighted risk factors. A person is given 5 points if identified as a gang member, if on parole or probation, for each incident "involving a gun" over the previous two years, for each violent crime arrest over the previous two years. People are given 1 point for every "quality" police contact over the last two years, based on FIs, arrests, and other reports (ibid: p.11).

This gives each individual a Chronic Offender Score and crime analysts use this method to create the required 12 Chronic Offender Bulletins, then rank order the bulletins based on points. Individuals with the

most points become the primary targets of patrol and special units (ibid).

Types of crime data processed by LASER (From SLSC interview): Automatic License Plate readers to track people; Past incarceration data including parole and probation data; Data from gang database; Arrest records; CCTV, Risk assessments; ArcGIS mapping system

The SLSC found LASER being used in relation to: "property crime, gun violence, whatever the division felt like they needed to use it for" (Garcia, 2020). She also said it was significant that LASER was rolled out to other divisions, at least 14 of the 21 LAPD divisions (SLSC, 2018: p.13).

Impact

The Stop LAPD Spying Coalition did a lot of work to find out how the systems worked and to access documents about these systems.

To our knowledge, no impact assessment was done before implementing PredPol and LASER.

Both systems were praised (Lopez, 2014) early on for causing an apparent reduction in crime. For instance both PredPol and the LAPD credited the PredPol system with helping bring about substantial reductions in property crime in the Foothill Division in 2012 through 2014, but crime crept back up after that (Jouvenal, 2016).

Ultimately the OIG's review found that "overall, targeted crime has decreased as police presence has increased; however, results broken down by quarter and area were more mixed. In general, given the difficulty of isolating the impact of these programs, as opposed to other factors that may impact crime, the OIG cautions against drawing strong conclusions from the available statistics (OIG, 2019: p.2).

"It didn't seem that there was a lot of contact that was driven by the program [nor were there a significant number] of officers using the program. So ultimately [...] our main finding was that there was so much inconsistency in how it was being used that it was very difficult to track[,] either the effectiveness in terms of preventing or addressing crime, and in terms of community concerns, of people's lives being impacted fairly or unfairly. So it was [...] kind of [an] inconclusive result." (Naguib interview, 2020)

Concerns were raised about the limits of the data being collected about how the systems were being used, which made reviewing them near impossible. Other concerns were raised about how the systems were not being implemented consistently.

Garcia (2020) says South LA was hit hard in 2016 by the roll-out LASER. She said "South LA is a predominantly Black area and for this area the rollout was pretty intense because they rolled out LASER in combination with setting up a data analytics centre in South Central called a Community Safety Operation Centre" (CSOC). This gave more agencies access to LASER data because it was shared through the CSOC. For example Mission Sheets would be given to Metro (a special enforcement arm of LAPD that is not committed to any particular division, and is effectively a "strong arm") and the Gang Enforcement Detail. These are all under the Major Crime Division's Special Operation Bureau, Counter Terrorism Division of LAPD. So that's who was receiving their mission from CSOC."

The Coalition's 2018 report was based on talking to impacted local communities about the effects of predictive policing. They explain that "most people offered stories of hyper police presence in certain communities, segregated communities, and profiling" (SLSC, 2018: p.35).

Concerns were raised that unhoused people were being aggressively targeted. The Coalition also raised concerns about plans to introduce real-time location-based alerts of people who were identified as 'Chronic Offenders.' There were concerns about this leading to the targeting and tracing of some of the most vulnerable community members.

Going forward

There is interest in continued development of the use of data driven systems. It has also been noted that the LAPD is convening a new oversight panel called Reimagining Public Safety that will be made up of 12 "expert" community members to provide oversight of surveillance.

Sources

Interviewees conducted online 2020

Jamie Garcia - Stop LAPD Spying Coalition

Hamid Khan - Stop LAPD Spying Coalition

Camelia Naguib - Assistant Inspector General, LAPD Office of the Inspector General

References

Centre for Evidence-Based Crime Policy (2015) EBP Hall of Fame: Sean Malinowski. Available at: https://cebcp.org/hall-of-fame/sean-malinowski/

Chang, C. (2018) *LAPD Officials Defend Predictive Policing As Activists Call For Its End*, *Los Angeles Times*. Available at: https://www.latimes.com/local/lanow/la-me-lapd-data-policing-20180724-story.html (Accessed: 26 August 2021).

Ferguson, A. G. (2016) *Policing Predictive Policing*. SSRN Scholarly Paper ID 2765525. Rochester, NY: Social Science Research Network. Available at: https://papers.ssrn.com/abstract=2765525 (Accessed: 26 August 2021).

Garcia, J. (2020) SLSC timeline of events for PredPol and LASER. Available at: https://docs.google.com/document/d/1XTRyWOVwlIB5pQTEyDvGmBIEtK6VNE1aZdCfyQwr9aE/edit

George Mason University (no date) *Sean Malinowski*, *The Center for Evidence-Based Crime Policy (CEBCP)*. Available at: https://cebcp.org/hall-of-fame/sean-malinowski/ (Accessed: 26 August 2021).

Haskins, C. (2020) *The Tool Was Supposed To Predict Crime. Now Los Angeles Police Say They Are Dumping It.*, *BuzzFeed News*. Available at: https://www.buzzfeednews.com/article/carolinehaskins1/los-angeles-police-department-dumping-predpol-predictive (Accessed: 26 August 2021).

Jouvenal, J. (2016) *Police Are Using Software To Predict Crime. Is It A 'Holy Grail' Or Biased Against Minorities?*, *The Washington Post*. Available at: https://www.washingtonpost.com/local/public-safety/police-are-using-software-to-predict-crime-is-it-a-holy-grail-or-biased-against-minorities/2016/11/17/525a6649-0472-440a-aae1-b283aa8e5de8_story.html (Accessed: 26 August 2021).

Lopez, L. (2014) *Cops Credit "Predictive Policing" for Zero Crime Day*, *NBC Los Angeles*. Available at: https://www.nbclosangeles.com/news/national-international/cops-credit-predictive-policing-zero-crime-day/75453/ (Accessed: 26 August 2021).

Los Angeles Police Commission meetings archive, https://www.lapdonline.org/police_commission_meetings_archives (Los Angeles Police Department, no date)

Los Angeles Police Department (2009) Smart Policing in Los Angeles: Targeting Gun Violence. Stop LAPD Spying Coalition. Available

at: https://stoplapdspying.org/wp-content/uploads/2018/05/Attachment-2-Program-Narrative.pdf

Los Angeles Police Commission (2018) *Los Angeles Police Commission Meeting*. Available at: https://www.youtube.com/watch?v=1g4n9mFQIfQ (Accessed: 26 August 2021).

Los Angeles Police Commission (2019) *Review Of Selected Los Angeles Police Department Data-Driven Policing Strategies*, *Google Docs*. Available at: https://drive.google.com/file/d/1lywhQK6TpohmBSanqBtyc1V9nWQERq9k/view?usp=sharing&usp=embed_facebook (Accessed: 26 August 2021).

Los Angeles Police Department (no date a) *Police Commission Meetings Archives*, *The Los Angeles Police Department*. Available at: https://www.lapdonline.org/police_commission_meetings_archives (Accessed: 26 August 2021).

Los Angeles Police Department (no date b) *The Function and Role of the Board of Police Commissioners*, *The Los Angeles Police Department*. Available at: https://www.lapdonline.org/police_commission/content_basic_view/900 (Accessed: 26 August 2021).

Los Angeles Police Department and Uchida, Dr. C. D. (2014) 'Smart Policing: Institutionalizing Operation LASER in The Los Angeles Police Department'. Available at: https://stoplapdspying.org/wp-content/uploads/2018/05/Program-Narrative.pdf.

Moravec, E. R. (2019) *Do Algorithms Have a Place in Policing?*, *The Atlantic*. Available at: https://www.theatlantic.com/politics/archive/2019/09/do-algorithms-have-place-policing/596851/ (Accessed: 26 August 2021).

PredPol (no date) *PredPol*, *PredPol*. Available at: https://www.predpol.com/ (Accessed: 26 August 2021).

Puente, M. (2019a) *LAPD Ends Another Data-Driven Crime Program Touted to Target Violent Offenders*, *Los Angeles Times*. Available at: https://www.latimes.com/local/lanow/la-me-laser-lapd-crime-data-program-20190412-story.html (Accessed: 26 August 2021).

Puente, M. (2019b) *LAPD To Change Crime Data Program As Activists Tell Police Commission To 'Shut It Down'*, *Los Angeles Times*. Available at: https://www.latimes.com/local/lanow/la-me-ln-lapd-commission-data-program-controversy-20190312-story.html (Accessed: 26 August 2021).

SLSC Timeline Of Events For PredPol And LASER (2020) *Stop LAPD Spying Coalition*. Available at: https://docs.google.com/document/d/1XTRyWOVwIIB5pQTEyDvGmBIEtK6VNE1aZdCfyQwr9aE/edit?usp=sharing&usp=embed_facebook (Accessed: 26 August 2021).

Smith, M. (2019) Intradepartmental Correspondence: Review of Selected Los Angeles Police Department Data-Driven Policing Strategies, Inspector General, Police Commission. Available at: https://www.muckrock.com/news/archives/2019/mar/12/algorithms-lapd-predpol/

Stop LAPD Spying Coalition (2018) 'Before the Bullet Hits the Body: Dismantling Predictive Policing in Los Angeles'. Stop LAPD Spying Coalition. Available at: https://stoplapdspying.org/wp-content/uploads/2018/05/Before-the-Bullet-Hits-the-Body-May-8-2018.pdf.

Stop LAPD Spying Coalition (no date) 'Smart Policing in Los Angeles: Targeting Gun Violence'. Stop LAPD Spying Coalition. Available at: https://stoplapdspying.org/wp-content/uploads/2018/05/Attachment-2-Program-Narrative.pdf.

Stop LAPD Spying Coalition Timeline of events (2020) (*SLSC Timeline Of Events For PredPol And LASER*, 2020)

Uchida, Craig et al (2014) "Smart Policing: Institutionalizing Operation LASER in The Los Angeles Police Department" Stope LAPD Spying Coalition. Available at: https://stoplapdspying.org/wp-content/uploads/2018/05/Program-Narrative.pdf

New Zealand Police Decide Not to Pursue Use of Clearview AI Facial Recognition Technology

Summary

New Zealand's High Tech Crime Group decided to test a trial copy of Clearview's Facial Recognition system. Some testing took place with some images uploaded to the system between January and March 2020. The public and it seems other government officials learned of the trial following reporting by RNZ.

Why was it not pursued?

It was reported that the test led to a disappointing match rate and so the High Tech Crime Group decided not to pursue the technology. After media coverage of the trial the Police Commissioner ordered a stocktake of uses of surveillance technologies.

Key factors leading to change

The system was quietly tried and then not pursued because it was not deemed effective, the match rate was found inadequate.

The public learned the system had been tried through media coverage after the decision had been made not to pursue this system. This coverage and public outcry led to controversy and concern about the use of the system violating people's privacy. It was noted in an interview with reporter Mackenzie Smith that there is concern in NZ about facial recognition technology (Smith interview, 2020). Controversy surrounded a lack of consultation with higher up officials (Smith, 2020a and 2020b).

How did the system work?

It has been reported that Clearview AI have claimed their system uses "a database of more than three billion images ... scraped from Facebook, YouTube, Venmo and millions of other websites" (Hill, 2020).

The system was tried in NZ between Jan and March 2020.

The technology was pursued by New Zealand Police's High Tech Crime Group after reading about the system in a New York Times article (Hill, 2020). In an interview, Chen said that this unit is responsible for using technology to solve crimes but tries to solve crimes which are committed using technology, so they have a broad remit that meant that looking at facial recognition technology seemed in their remit.

The High Tech Crime Group wanted to see if the technology was worth pursuing and contacted Clearview without notifying superiors (Chen interview, 2020).

Reports suggest that some testing took place with test images uploaded to the trial copy of Clearview AI's system.

Impact

No impact assessment was carried out. There was some discussion with the chief privacy officer for the police, internally, who suggested the Privacy Commissioner be consulted if the system was pursued.

Going forward

In August 2020 it was reported (Pennington, 2020) that police in New Zealand were setting up a facial recognition system to identify people from CCTV feeds.

Police use of facial recognition technologies is in the news (Pennington, 2020) again in September and October as it has been reported that the Department of Internal Affairs has signed a master agreement with DXC Technology which enables public and private pursuit of uses of facial recognition.

Sources

Media: RNZ (Mackenzie, 2020), RNZ (Mackenzie, 2020), TechWire Asia (Devanesan, 2020), NZ Herald (Mackenzie, 2020)

Interviews conducted online, 2020:

Andrew Chen (University of Auckland)

Mackenzie Smith (Radio New Zealand)

References

Devanesan, J. (2020) *New Zealand Police Trial Controversial Clearview AI Facial Recognition Tech*, *Tech Wire Asia*. Available at: https://techwireasia.com/2020/05/police-trial-controversial-clearview-ai-facial-recognition-tech-in-new-zealand/ (Accessed: 13 August 2021).

Gavaghan, C. *et al.* (2020) *On Facial Recognition, Police, and Regulation of New Technologies*, *Koi Tū: The Centre for Informed Futures*. Available at: https://informedfutures.org/on-facial-recognition-police-and-regulation-of-new-technologies/ (Accessed: 31 August 2021).

Heilweil, R. (2020) *The World's Scariest Facial Recognition Company, Explained*, *Vox*. Available at: https://www.vox.com/recode/2020/2/11/21131991/clearview-ai-facial-recognition-database-law-enforcement (Accessed: 31 August 2021).

Hill, K. (2020a) 'Facial Recognition Start-Up Mounts a First Amendment Defense', *The New York Times*, 11 August. Available at: https://www.nytimes.com/2020/08/11/technology/clearview-floyd-abrams.html (Accessed: 31 August 2021).

Hill, K. (2020b) 'The Secretive Company That Might End Privacy as We Know It', *The New York Times*, 18 January. Available at: https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html (Accessed: 31 August 2021).

Mac, R. (2020) *Clearview AI Says It Will No Longer Provide Facial Recognition To Private Companies*, *BuzzFeed News*. Available at: https://www.buzzfeednews.com/article/ryanmac/clearview-ai-no-facial-recognition-private-companies (Accessed: 31 August 2021).

Mac, R., Haskins, C. and McDonald, L. (2020) *Clearview's Facial Recognition App Has Been Used By The Justice Department, ICE, Macy's, Walmart, And The NBA*, *BuzzFeed News*. Available at: https://www.buzzfeednews.com/article/ryanmac/clearview-ai-fbi-ice-global-law-enforcement (Accessed: 8 July 2021).

Mackenzie, S. (2020a) *Police 'Stocktake' Surveillance Tech After Clearview AI Facial Recognition Trial*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/416913/police-stocktake-surveillance-tech-after-clearview-ai-facial-recognition-trial (Accessed: 13 August 2021).

Mackenzie, S. (2020b) *Police Trialled Facial Recognition Tech Without Clearance*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/416483/police-trialled-facial-recognition-tech-without-clearance (Accessed: 13 August 2021).

Merlan, A. (2020) *Here's the File Clearview AI Has Been Keeping on Me, and Probably on You Too*. Available at: https://www.vice.com/en/article/5dmkyq/heres-the-file-clearview-ai-has-been-keeping-on-me-and-probably-on-you-too (Accessed: 31 August 2021).

Pennington, P. (2019) *Police Open to Using Facial Recognition From Auckland Transport CCTV Cameras*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/396716/police-open-to-using-facial-recognition-from-auckland-transport-cctv-cameras (Accessed: 31 August 2021).

Pennington, P. (2020a) *Government Facial Recognition Tech Deal Offers Wide Access*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/428103/government-facial-recognition-tech-deal-offers-wide-access (Accessed: 13 August 2021).

Pennington, P. (2020b) *Police Setting Up $9m Facial Recognition System Which Can Identify People From CCTV Feed*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/424845/police-setting-up-9m-facial-recognition-system-which-can-identify-people-from-cctv-feed (Accessed: 13 August 2021).

RNZ (2020) *Police Ending Armed Response Teams After Trial - Commissioner*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/418593/police-ending-armed-response-teams-after-trial-commissioner (Accessed: 31 August 2021).

Roy, E. A. (2020) *New Zealand Politician Admits Leaking Covid-19 Private Details to Media*, *the Guardian*. Available at: http://www.theguardian.com/world/2020/jul/07/new-zealand-politician-admits-leaking-covid-19-private-details-to-media-hamish-walker (Accessed: 31 August 2021).

Smith, M. (2020) *Police Searched for Suspects in Unapproved Trial of Facial Recognition Tech, Clearview AI*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/416697/police-searched-for-suspects-in-unapproved-trial-of-facial-recognition-tech-clearview-ai (Accessed: 31 August 2021).

Woollacott, E. (2020) *UK And Australia To Probe Clearview AI Facial Recognition*, *Forbes*. Available at: https://www.forbes.com/sites/emmawoollacott/2020/07/09/uk-and-australia-to-probe-clearview-ai-facial-recognition/ (Accessed: 31 August 2021).

Durham Police cancel HART (Harm Risk Assessment Tool)

Summary

One of the first algorithmic risk assessment tools to be used by a UK police force involved the development of HART. The system was started in 2013 and it was implemented in mid-2016. HART was the result of a collaboration between Durham Police and statistical experts at Cambridge University and formed part of an ongoing partnership between the two institutions. HART was created to help officers decide which individuals could be referred to Durham Police's rehabilitation program, Checkpoint, designed to find alternatives to prosecution. Problems with accuracy were found in a 2016 validation study of the model, in that the tool was better at predicting low risk offenders than high risk offenders. In order to compensate for the possibility of wrongly predicting a high-risk offender to be low risk the HART algorithm overestimated high risk predictions.

HART also controversially made use of Experian's Mosaic segmentation tool, which was revealed in a 2018 investigation by Big Brother Watch. One of the HART project leaders said the force stopped using Mosaic data in 2018.

Why was it stopped?

Durham's Freedom of Information department confirmed that HART was dropped in September 2020. We do not know why they stopped using HART.

Key factors leading to change?

Acceptance of HART, as outlined in reports and publications, was dependent on performance. It was outlined in multiple places that: "Durham's test of HART is a crucial first step to determining what place, if any, algorithmic forecasting techniques have in policing" (Oswald et al, 2017: p.231).

The use of the system and its effectiveness has involved ongoing investigation. Research suggests the system is not in use, but we do not know if this means it will not be developed further and used in future and we do not know why use has stopped. Previous research indicated a 24-month period of use to evaluate its predictions and assessment of its use after that time (Hatterslea 2018).

How did the system work?

HART was a risk assessment tool (HART stands for Harm Risk Assessment tool) that was used to predict the likelihood of recidivism among offenders and was the first predictive policing tool to be used in decision making about individuals (Big Brother Watch, 2018a).

As detailed by Oswald et al: "the central goal of the development team was to promote consistency in decision making, enabling targeted interventions and rigorous testing to find responses to offending that reduce future harm and recidivism" (2018: p.227). It has been noted that cuts to police budgets placed pressure to find systems to triage offenders and it was determined that the use of effective forecasting could help (Oswald, 2018: p.231).

HART was initially tested in 2013 (Burgess, 2018; Babuta, 2017: p.23) and it was implemented in mid-2016 (Cambridge University, 2018). HART was the result of a collaboration between Durham Constabulary and statistical experts at Cambridge University and formed part of an ongoing partnership between the two

institutions (Nilsson, 2019; Oswald et al: p. 227).

The system was used to sort individuals in police custody into three categories of predicted recidivism risk: high, medium and low. High Risk indicated an individual was likely to commit a new serious offence over the next two years, Medium Risk was given for those predicted to commit a non-serious crime over the next two years, and Low Risk was assigned to those not expected to commit a crime over the next two years (Oswald et al, 2018: p.227).

HART was created to help officers decide which individuals could be triaged/referred to Durham's Operation Checkpoint program, a "culture-changing" rehabilitation initiative designed to "tackle the root causes of offending and associated health and community issues by offering an alternative to prosecution for a very specific subset of criminal offenders" (Oswald et al, 2018: p.227). Checkpoint is important to understanding the development of HART because it prefaced the system and was designed to address the 'revolving door' in the criminal justice system, i.e. "low-level offenders cycling in and out of the system" (RSA, 2019: p.39). According to the HART project lead Sheena Urwin, the use of HART within Checkpoint helps to better target those categorized as low-level offenders and present them with an alternative to prosecution through a formal four-month contract (2019). As Burgess (2018) highlights, "the scheme is designed to intervene in proceedings rather than push people through the UK's court system." Only individuals categorised as Medium Risk by HART qualified for the Checkpoint program (Oswald et al:), and it was hoped that feeding HART into Checkpoint would help "turn moderate risks into low risks" (Cambridge University, 2018).

The HART algorithm was developed using a form of machine learning called random forests2, said to be desirable for its "ability to detect relatively rare but dangerous outcomes, to model relationships in non-linear ways, and to balance the differential costs of different kinds of errors" (Oswald 2018). The HART model was built using historical data, approximately 104,000 custody events over a five-year period (2008-2012), and did not include any data held on the Police National Computer or Police National Database (Cambridge University?: p.228), so in this sense was a "very local" dataset (Babuta, 2017: p.24). HART used 34 different predictors to arrive at a forecast, 29 of which focused upon the offender's prior history of criminal behaviour (Babuta: p.228). These 29 variables were combined with age, gender, two forms of residential postcode, as well as the count of existing police intelligence reports relating to the offender, and the combination of these variables produced the output risk level (ibid). Race was intentionally not used as a predictor variable (Burgess, 2018). All conclusions reached by HART were based on 509 votes by the system and a vote was either low, moderate or high (Burgess, 2018). With avoidance of negative feedback loops in mind, the primary postcode predictor was limited to the first four characters of the postcode, which usually designated a sizable geographic area (2018).

The random forest algorithm was designed to treat errors in overprediction (i.e. overestimates of an individual's likelihood of reoffending) as less serious and less "costly" than errors in underprediction or underestimates of risk, which were seen as dangerous to public safety. This means the model was intentionally weighted towards overprediction errors, with the dangerous underestimate errors occurring less frequently, with a ration set so that the model produced roughly two cautious errors (overestimate) for each dangerous error (underestimate) (ibid). One of the HART project leaders, Sheena Urwin (who is Head of Criminal Justice at Durham Constabulary and also a graduate of Cambridge University's Institute of Criminology's Police Executive Master of Studies Programme), is quoted as saying "not all errors are created equal", by which she referred to the idea that some forms of inaccuracy have worse consequences than others.

According to Oswald et al. (2018) HART was meant to support and inform human decision-making, not replace it, partly due to the perceived limitation that the system only used data held within Durham Constabulary's databases so did not have "all the available information": "With both their own local knowledge and their access to other data systems, custody officers will frequently be aware of other information that overrides the model's predictions, and they must apply their own judgement in deciding upon the disposition of each offender's case" (Oswald et al., 2018: p.230).

Impact

In response to a 2016 independent validation study, Durham Constabulary did further work on ethical issues. It undertook awareness sessions relating to unconscious bias and utilising HART as a discussion topic, with some sessions aimed specifically at custody police officers to ensure officers understood HART and viewed it is a decision support tool "that cannot know all the information available to a human being" (Oswald et al, 2018: p.230). According to Burgess these sessions aimed "to prevent existing human biases – around race and social prejudices – creeping into use with HART" (2018).

From Oswald et al's review of HART it is clear that the project team were well aware of the potential for bias, for example in reflecting on Mittlestadt et al.'s five ethical concerns posed by algorithms the review's authors state that "the implementation of the HART model raises every single one of these concerns to a greater or lesser extent" (2018: p.232). These concerns include the potential for discriminatory outcomes and problems with evidence having the potential to lead to opacity or bias (2018: p.231). Particular attention was given in the review to the inclusion of postcode data in the HART model, which the authors identified as risking the emergence of a negative feedback loop "that may perpetuate or amplify existing patterns of offending" (2018: p.230). As detailed in the review: "If the police respond to forecasts by targeting their efforts on the highest-risk postcode areas, then more people from these areas will come to police attention and be arrested than those living in lower-risk, untargeted neighbourhoods. These arrests then become outcomes that are used to generate later iterations of the same model, leading to an ever-deepening cycle of increased police attention." (Oswald et al., 2018). More research is needed to determine how these concerns about bias were addressed in practice.

Concerns were raised by Big Brother Watch in April 2018, when the NGO found that Durham Police had been using Experian's Mosaic Public Sector dataset (BBW, 2018a: 2018b). BBW also raised concerns about the use of postcode data in HART, saying this was problematic in AI tools "because [it] carries a risk of perpetuating bias towards areas marked by community deprivation" (BBW 2018a). BBW also commented that the use of a segmentation tool like Mosaic "risks incredible prejudice in our justice system" and questioned whether it was right to use such data even if it is collected by legal means (2018a).

HART was tweaked to no longer include postcode data in 2018. For instance, Oswald et al's (2018) review of HART stated that "HART is currently being refreshed with more recent data, and with an aim of removing one of the two postcode predictors" (Oswald, 2017: p.230), but no reference is made in this paper to Experian or Mosaic. According to Burgess the primary postcode predictor, which included the first four digits, was the variable that was removed (Burgess, 2018).

A comparison of police officer decisions and the algorithmic system was conducted. Use of the first live version of HART required that Durham custody officers make their own predictions of each offender's future arrests whenever the algorithm was used, allowing the Constabulary to directly compare HART's predictions with the human judgement of officers (Oswald et al, 2018: p.233). At the time of Oswald et al's review of HART, initial results showed that custody officers were "generally uneasy with forecasting at either extreme and avoided making both high and low risk predictions" (2018). A significant proportion of

officer predictions were for moderate risk behaviour (63.5%), and the model and officers agreed only 56.2% of the time.

Further work by Durham Police in relation to considerations of impact was their collaboration with Marion Oswald was the development of a decision-making guidance framework for the deployment of algorithms in policing contexts - based on Durham's experience of HART - called 'Algorithms in Policing –Take ALGO-CARE™' (Oswald et al, 2018). This framework was intended to translate key public law and human rights principles into practical considerations and guidance that can be addressed by public sector bodies (Oswald and Urwin, 2017). Its creators have said that unlike other auditing and ethical frameworks which utilise high-level principles, Algo-care is intended to provide practical guidance that practitioners can refer to in their day-to-day work (ibid). Interestingly, the framework is designed to enable controlled algorithmic experimentation in the public sector whilst "carefully managing any risks to individual rights" (Oswald et al, 2018: p.227).

In 2019 HART was one of the case studies presented to citizen juries as part of the Royal Society of Arts exploration of AI in decision making. The RSA reports that during deliberation citizens probed the experts on the data used by HART to calculate risk and found there was some concern about whether the use of postcode data may result in biased outcomes, with citizens questioning what safeguards were in place to prevent bias (RSA, 2019: p.40). Some citizens also expressed concern about HART's accuracy rate of 62 percent and discussed whether or not this was an acceptable threshold for use (RSA, 2019).

Sources

Media: Wired (Burgess, 2018), Financial Times (Nilsson, 2019)

Civil society: Big Brother Watch (Big Brother Watch, 2018)

Academia: Cambridge University (University of Cambridge, 2018), Oswald et al review of HART (Oswald et al., 2017)

References

Babuta, A (2017). "Big Data and Policing: An Assessment of Law Enforcement Requirements, Expectations and Priorities." Occasional paper for Royal United Services Institute (RUSI). Available at: https://rusi.org/sites/default/files/201709_rusi_big_data_and_policing_babuta_web.pdf

Barton, M. (2017) Centre for Evidence-Based Crime Policy (CEBCP) (2017) EBP Hall of Fame: Michael Barton. https://cebcp.org/hall-of-fame/michael-barton/

BBC News (2018) 'Durham Police Criticised over "Crude" Profiling', *BBC News*, 9 April. Available at: https://www.bbc.com/news/technology-43428266 (Accessed: 8 September 2021).

Big Brother Watch (2018a) *A Closer Look at Experian Big Data and Artificial Intelligence in Durham Police*, *Big Brother Watch*. Available at: https://bigbrotherwatch.org.uk/2018/04/a-closer-look-at-experian-big-data-and-artificial-intelligence-in-durham-police/#_ftn2 (Accessed: 15 July 2021).

Big Brother Watch (2018b) *Police use Experian Marketing Data for AI Custody Decisions — Big Brother Watch*, *Big Brother Watch*. Available at: https://bigbrotherwatch.org.uk/2018/04/police-use-experian-marketing-data-for-ai-custody-decisions/ (Accessed: 8 September 2021).

Bridge, M. (2018) 'Police Computer Locks up Suspects Based on Postcode', *The Times*, 9 March. Available at: https://www.thetimes.co.uk/article/police-computer-locks-up-suspects-based-on-postcode-dz0qknzrr (Accessed: 8 September 2021).

Burgess, M. (2018) *UK Police are Using AI to Make Custodial Decisions – But It Could Be Discriminating Against the Poor*, *WIRED UK*. Available at: https://www.wired.co.uk/article/police-ai-uk-durham-hart-checkpoint-algorithm-edit (Accessed: 15 July 2021).

Cambridge University (2018). 'Helping police make custody decisions using artificial intelligence.' (University of Cambridge, 2018), Cambridge University research blog. 26 February. https://www.cam.ac.uk/research/features/helping-police-make-custody-decisions-using-artificial-intelligence

Data Justice Lab (2018) Freedom of Information request response 422-18. https://data-scores.org/920f7229f31040eacdd410f7778410142415923d129da615113ff20af6e62854 (Data Scores in the UK, 2018)

Data Scores in the UK (2018) [Data Justice Lab request response] FOI 422-18 Response .docx, Data Scores in the UK. Available at: https://data-scores.org/920f7229f31040eacdd410f7778410142415923d129da615113ff20af6e62854 (Accessed: 8 September 2021).

Durham Constabulary (2020). "Police staff member wins national accolade." August 13.

https://www.durham.police.uk/news-and-events/Pages/News%20Articles/Police-staff-member-wins-national-accolade.aspx [page not found]

Durham Constabulary (2015). Checkpoint information and advice page. https://www.durham.police.uk/Information-and-advice/Pages/Checkpoint.aspx [Accessed October 26 2020]

George Mason University (no date) *Michael Barton*, *The Center for Evidence-Based Crime Policy (CEBCP)*. Available at: https://cebcp.org/hall-of-fame/michael-barton/ (Accessed: 8 September 2021).

Hattersley, A (2018). Durham Constabulary response to FOI request. Information rights and disclosure unit. Data Justice Lab. April 25. https://data-scores.org/2cd66b135f83a140b68afa048ead51207772a530c83b322a4efc1f865b4017d1

Nilsson, P. (2019) 'UK Police Test if Computer can Predict Criminal Behaviour', *Financial Times*, 6 February. Available at: https://www.ft.com/content/9559efbe-2958-11e9-a5ab-ff8ef2b976c7 (Accessed: 15 July 2021).

Oswald, M. *et al.* (2017) *Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and 'Experimental' Proportionality*. SSRN Scholarly Paper ID 3029345. Rochester: Social Science Research Network. doi:10.2139/ssrn.3029345.

Oswald, M and Urwin, S (2017). Written evidence submitted to Parliament's Science and Technology Committee for the review 'Algorithms in Decision-Making'. (ALG0030) UK Parliament. April 2017. http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee/algorithms-in-decisionmaking/written/69002.html (Accessed: 8 September 2021).

The Forum for Ethical AI (2019) 'Democratising Decisions About Technology: A Toolkit'. Royal Society for the Encouragement of Arts, Manufactures and Commerce.

The Howard League for Penal Reform (no date) *2016 Community Awards Winners*, *The Howard League for Penal Reform*. Available at: https://howardleague.org/community-awards/2016-community-awards-winners/ (Accessed: 8 September 2021).

University of Cambridge (2018) *Helping Police Make Custody Decisions Using Artificial Intelligence*, *University of Cambridge*. Available at: https://www.cam.ac.uk/research/features/helping-police-make-custody-decisions-using-artificial-intelligence (Accessed: 15 July 2021).

Child Welfare ADS

Overview

The case study investigations of cancelled child welfare ADS involved the study of systems that were cancelled at the development and planning stage, after piloting and after implementation. These case studies also present differences in terms of how much information is publicly available about them, raising important questions about transparency and accountability.

Child welfare

Denmark	Denmark decides not to pursue use of the Gladsaxe model. (2017 - 2018)
United States, Illinois	Illinois Department of Children and Family Services (DCFS) stops use of Rapid Safety Feedback (RSF) (2015 - 2017)
New Zealand	Government decides not to use Predictive Risk Modelling to identify children at risk of abuse and neglect (2012 - 2015).
United Kingdom, Hackney	Hackney Council decides not to pursue use of Early Help Profiling System (2015 - 2019).

Replicating inequality

The Gladsaxe model proposed for Denmark and the predictive risk model considered for use in New Zealand were both developed to identify children in need of help early for the purposes of early intervention. When information about the plans for these systems became publicly known they were criticized by researchers and politicians. A number of critiques about the system are similar to critiques raised about predictive policing systems in Los Angeles as well as the fraud detection systems referenced in our case studies. These critiques centre on the potential for ADS to further embed bias and lead to greater inequality, in addition to concerns about accuracy and bias. Unique in the case of New Zealand was that the predictive model was made public which enabled review and public debate.

Validity

In contrast to these two systems, little is known about why the Illinois Department of Family Services cancelled use of the Rapid Safety Feedback programme. As with some of the other case studies, reporting on this case suggests it was cancelled due to issues with *effectiveness*. In this case, there were concerns with accuracy and overprediction. In the case of the Early Help Profiling System used in the UK, a government spokesperson linked concerns about effectiveness to limitations of the data used in the system.

Denmark, Gladsaxe

Summary

The 'Gladsaxe model' was a classification system developed by the Gladsaxe municipality to trace «children who were vulnerable due to social circumstances before they presented as in need (Algorithm Watch 2019).

Gladsaxe municipality developed the model. Gladsaxe and two other municipalities (Guldborgsund and Ikast-Brande) requested exemption from legislation in October 2017 to use it. The government declines the exemption because they want to change the legislation to allow all municipalities to use this kind of system. There is critical media coverage in response to this and public and academic protest. In December 2018 there is also an unrelated data leak in Gladsaxe municipality. The government changes its mind and stops plans to change the legislation and Gladsaxe is not permitted to implement their model. The model has not been used.

The government planned to make it legal for all 98 municipalities to use this model to risk assess children. After a strong public reaction and criticism, the Liberal Alliance Government stated the proposal had been shelved in December 2018.

Why did plans change?

The system was criticized publicly, politically, and also by academic researchers. An unrelated data leak at the same time raised concerns about government use of confidential data. The Liberal Alliance's spokeswoman Christina Egelund stated that municipalities were not equipped to deal with «the great responsibility that lies in taking care of the personal data of the citizens».

Key factors leading to change

The system was criticized by the public, politicians and academic researchers. A news story in Politiken raised concerns as did think tank Justitia as well as academics. A data leak at the same time created greater public awareness around concerns about protection of sensitive data. In response to these factors the government said it had shelved the proposal.

How did the system work?

The model was developed in-house by the municipality. Developers said they identified a number of "significant edge risk indicators" that explain a substantial part of the variance in the well-being of children and young people. They wanted to design a data driven tracking model to identify risk indicators in parents before special needs symptoms are experienced by children.

Three municipalities wanted to collect and combine information from different public sources and to categorize it according to specific "risk indicators." The goal was to support the automatic detection of children with special needs earlier. The risk estimation included things like parental mental illness, unemployment, missing a doctor's appointment or dentist appointment and divorce. The risk estimation was to be applied to all children.

The aim of the model was to select families with children at risk of social vulnerability by pooling data from the field of dental and health, social affairs, employment, family data, education, and day care. The model looked for risk indicators for adults with children aged 0-6 years who have one or more entrances to the municipality. Reports suggest the model involved 9 different IT systems and 44 risk indicators.

The model was developed by using data on already tracked children who had reached 15. The model was described as a classification algorithm that determined a combination of risk indicators that would be most relevant to look at (Hovmand interview, 2020).

An unrelated event in this case is relevant as it influenced public debate about government use of personal data. In December 2018 a computer storing confidential information about 20,000 citizens in Gladsaxe Municipality was stolen from City Hall. Through an error, the spreadsheet storing this information was stored locally and so not protected through normal measures. In response to this the municipality's spokeswoman said: "I am not prepared to implement that proposal systematically before the municipalities are much better equipped to handle the great responsibility that lies in taking care of the citizens' personal information. There is no indication that the municipalities today are ready to handle more complex tasks with the interconnection of data» (Kjaer, 2020).

Another related event included the plan in Denmark to ensure no "ghettos" by 2030 (Seeman 2021). The early detection of vulnerable children was listed as part of the strategy, which was said to be inspired by the model plan put forward by Gladsaxe Municipality.

Impact

The system was not implemented.

The goal was to intervene in the lives of children and families earlier. This aim was fuelled by research indicating how important the early years are for children's development. The goal was also to prevent family break up.

Concerns were raised about plans to use citizen data without consent, rights, the use of a "points system" to evaluate families, concerns about the mass monitoring of citizens

Going forward

There remains interest in developing a data-based method to detect need and intervene earlier. Hovmand (2020) argues: "Don't we also have an ethical obligation to try to help them (families), if it's possible? This would be this double ethical dilemma that everything in this area is concerned with. That's why we wanted to do a small project that's externally reviewed and everything to get some experience to better be able to enable the politicians to decide if we should go this way or if we shouldn't."

In 2018 the Minister of Children and Social affairs noted that they want to find a model that "balances the diverse interests in the best way possible, so that the municipalities can act earlier with the help of vulnerable children and young people – while also ensuring a clean and safe framework where citizens' legal security is given high priority. It requires thorough preparation. At the same time, I have also said that I want to discuss a model with other parties so that it is not rushed through" (Mercado quoted in Politiken, English translation; Kjær 2018b).

Interviews conducted online in 2020:

Thomas Berlin Hovmand (Director of education and culture in the municipality of Gladsaxe also responsible for digital transformation in the municipality).

References

Aarhus University (no date) *TrygFonden's Centre for Child Research*, *TrygFonden's Centre for Child Research*. Available at: https://

childresearch.au.dk/en/ (Accessed: 8 September 2021).

Algorithm Watch (2019) 'Automating Society', *AlgorithmWatch*. Available at: https://algorithmwatch.org/en/automating-society/ (Accessed: 18 August 2021).

Andersen, T. (2019) *Gladsaxe taler ud om dataovervågning: Vi vil gerne lave en 'black box'*, *Version2*. Available at: https://www.version2.dk/artikel/gladsaxe-taler-ud-dataovervaagning-vi-vil-gerne-lave-black-box-1087552 (Accessed: 8 September 2021).

Brandimarte, E. (2019) *Should We Let the Algorithms Decide? A Critical Assessment of Automated Decision-Making in Europe*. College of Europe.

Datatilsynet (2020) *To kommuner indstillet til bøde*, *Datatilsynet*. Available at: https://www.datatilsynet.dk/presse-og-nyheder/nyhedsarkiv/2020/mar/to-kommuner-indstillet-til-boede (Accessed: 8 September 2021).

Kjær, J.S. (2018a) *Kommune advarer: Fortrolige oplysninger om 20.000 borgere er faldet i hænderne på kriminelle*, *Politiken*. Available at: https://politiken.dk/indland/art6910988/Kommune-advarer-Fortrolige-oplysninger-om-20.000-borgere-er-faldet-i-h%C3%A6nderne-p%C3%A5-kriminelle (Accessed: 8 September 2021).

Kjær, J.S. (2018b) *Regeringen har lagt sin plan om overvågning af børnefamilier i skuffen*, *Politiken*. Available at: https://politiken.dk/indland/art6919255/Regeringen-har-lagt-sin-plan-om-overv%C3%A5gning-af-b%C3%B8rnefamilier-i-skuffen (Accessed: 8 September 2021).

Kjær, J.S. (2018c) *Regeringen vil overvåge alle landets børnefamilier og uddele point*, *Politiken*. Available at: https://politiken.dk/indland/art6365403/Regeringen-vil-overv%C3%A5ge-alle-landets-b%C3%B8rnefamilier-og-uddele-point (Accessed: 8 September 2021).

Regeringen.dk (2018a) *Regeringen vil gøre op med parallelsamfund*, *Regeringen.dk*. Available at: https://www.regeringen.dk/nyheder/2018/ghettoudspil/ (Accessed: 8 September 2021).

Regeringen.dk (2018b) *Regeringen vil sikre tidlig opsporing af udsatte børn*, *Regeringen.dk*. Available at: https://www.regeringen.dk/nyheder/2018/ghettoudspil/tidlig-opsporing-af-udsatte-boern/ (Accessed: 8 September 2021).

Seemann A. The Danish 'ghetto initiatives' and the changing nature of social citizenship, 2004–2018. *Critical Social Policy*. 2021;41(4):586-605.

Trygfonden (no date) 'Invitation til Deltagelse'. Aarhus University.

Illinois Department of Family Services abandoned predictive analytics programme Rapid Safety Feedback

Summary

In 2015 Illinois Department of Children and Family Services (DCFS) implemented Rapid Safety Feedback (RSF), a predictive analytics tool developed by the non-profit Eckerd Connects and its for-profit partner Mindshare Technology. It was brought in by former DCFS Director, George Sheldon, who was hired to address Illinois' increase in child deaths. The implementation of RSF was central to his reform plans.

Why was it stopped?

Unreliability, inaccuracy and overprediction which overloaded caseworkers with new cases; issues with the contract including non-transparent bidding; the stark language the system uses (e.g. predicting likelihood of "death") was also cited as alarming for child welfare agencies. It was dropped in December 2017.

Key factors leading to change

Ultimately RSF was cancelled because the new DCFS Director, Beverly Walker, decided not to renew the contract. Walker is quoted in the Chicago Tribune as saying the "predictive analytics [wasn't] predicting any of the bad cases" (Jackson and Marx, 2017).

John Kelly observed in the Imprint (a youth and family news outlet) that this decision appeared to rest on three factors: caseworkers being alarmed by the language and stark (over)predictions of RSF (that was originally reported by the Chicago Tribune reporters who spoke to caseworkers); the fact that RSF failed to predict two controversial child deaths; and the fact that the contract was awarded on a no-bid basis, the controversy of which contributed towards Sheldon's resignation in on May 31 2017 (Kelly, 2017).

There was also controversy around how the original contract was awarded, made public through an internal audit. In 2017 Illinois Office of the Executive Inspector General and the then Director of Illinois' Department of Children and Family Services (DCFS), George Sheldon. An audit summary report shows that in April 2017 the investigation found that the DCFS wrongly processed the Eckerd no-bid contract as a grant instead of a sole source procurement or competitively bid contract, along with other findings of misconduct (OEIG, 2017b; OEIG 2017a: p.37).

Critical media coverage in the Chicago Tribune raised concerns about the awarding of the contract as well as how effective the system was (Jackson and Mark, May and Dec. 2017).

How did it work?

The RSF contract was awarded in September 2015 and was terminated/not renewed in December 2017.

The RSF system was originally developed by Eckerd Connects (based in Florida and formerly Eckerd Kids).

There were three phases to the implementation of RSF for DCFS: development of the model, deployment of the model, and ongoing hosting and support. Eckerd sub-contracted the development and deployment of the predictive model to Mindshare as well as the ongoing maintenance and support, while Eckerd itself

provided project management, case selection criteria and critical investigative practice (Illinois DCFS contract, 2015: p.7). RSF reportedly cost $366,000 (Jackson and Marx, Dec 2017).

Quality assurance staff at Eckerd Connects, a national non-profit provider of multiple child welfare services, use the Rapid Safety Feedback (RSF) tool which Mindshare - its for-profit partner - has developed to produce "real-time data and agency performance dashboards" (Eckerd Connects, 2020). It works by analysing historical data to predict future risk (Jackson and Marx, Dec 2017b) and is a fully automated system (DCFS contract, 2015: p.6).

RSF interoperated with DCFS data systems and data sharing protocols were established in order for the Eckerd/Mindshare team to have access to the state of Illinois' case-tracking system (DCFS contract, 2015: p.6). An automated "data dump" happened on a nightly basis that transferred data from this case-tracking system to Mindshare databases (Jackson and Marx, Dec 2017). This data was then analysed in order to assign a score of 1 to 100 to all "incoming investigations" (DCFS, 2015: p.6), that is, every child that had been the subject of an abuse allegation via the agency hotline (Jackson and Marx, 2017b). The scores were ready for DCFS by the next morning and staff were able to review the cases through a web-based secure portal, which was also supposed to present staff with "review questions to be answered, the documentation and tracking of any follow up activities required of the investigator and data for analysis" (DCFS contract, 2015: p.6).

However an Eckerd spokesperson told the Chicago Tribune that front-line caseworkers should "never" receive the raw scores, let alone make decisions based on them; instead the data instead should be reviewed by DCFS supervisors who are trained and coached by Eckerd to decide which cases need immediate attention and how to tackle them (Jackson and Marx, 2017a).

According to DCFS public statements, RSF algorithms rated children's risk of being killed or severely injured during the next two years (Jackson and Marx, 2017b). This risk was determined by attributes that were highly correlated with serious harm in past abuse cases, such as parents' ages, previous criminal records, evidence of substance abuse, or the presence of a new girlfriend or boyfriend (Jackson and Marx, 2017b).

Impact

We were unable to find any information about an impact assessment or user testing.

Media reports indicate the system was not as effective as expected. Some of this is connected to the data relied upon. Interviews reported in the Chicago Tribune suggest that the DCFS automated case-tracking system that interoperated with Mindshare software was riddled with data entry errors in two cases of controversial child deaths (Semaj Crosby and Itachi Boyle) that happened while RSF was in use. In addition, these interviews revealed that the DCFS system did not link investigations about many children to cases regarding their siblings, or other adults in the same home (Jackson and Marx, Dec 2017).

Jackson and Marx report that "caseworkers were alarmed and overwhelmed by alerts as thousands of children were rated as needing urgent protection"; for example RSF predicted that 369 young children, all under age 9, had a "100 percent chance of death or serious injury" in the next two years (Dec 2017). Further, Illinois child care agencies told the Tribune they were alarmed by computer-generated alerts such as: «Please note that the two youngest children, ages 1 year and 4 years have been assigned a 99% probability by the Eckerd Rapid Safety Feedback metrics of serious harm or death in the next two years.» Eckerd later told the Tribune it regrets using stark language suggesting the company could predict the probability of harm to a child (Eckerd Connects 2017) «We all agree that we could have done a better job with that

language. I admit it is confusing,» said Eckerd spokesman Douglas Tobin.

Sources

We approached four people for interviews, two didn't respond and two declined.

Media articles referenced: Governing (Jackson and Marx, 2017) and Chicago Tribune (Jackson and Marx, 2017)

References

Eckerd Connects (no date) *Eckerd Rapid Safety Feedback*, *Eckerd Connects*. Available at: https://eckerd.org/family-children-services/ersf/ (Accessed: 9 September 2021).

Hickey, B.K. (2016) *Saving Children, One Algorithm at a Time*, *GCN*. Available at: https://gcn.com/articles/2016/07/26/child-welfare-analytics.aspx (Accessed: 9 September 2021).

Illinois Department of Children and Family Service (2015) 'Illinois Department of Children and Family Services Eckerd Rapid Safety Feedback'. MuckRock. Available at: https://www.muckrock.com/foi/illinois-168/illinois-department-of-children-and-family-services-eckerd-rapid-safety-feedback-31206/ (Accessed: 9 September 2021).

Illinois Department of Children and Family Services (2016) 'Illinois Child Welfare Transformation: 2016-2021 Transformation Plan'.

Jackson, D. and Marx, G. (2017a) *Data Mining Program Designed to Predict Child Abuse Proves Unreliable, DCFS says*, *Chicago Tribune*. Available at: https://www.chicagotribune.com/investigations/ct-dcfs-eckerd-met-20171206-story.html (Accessed: 29 June 2021).

Jackson, D. and Marx, G. (2017b) *Too Much Data? Illinois Abandons System Meant to Predict Child Abuse*, *Governing*. Available at: https://www.governing.com/archive/tns-chicago-data-mining.html (Accessed: 8 September 2021).

Jackson, D., Marx, G. and Eldeib, D. (2017) *Director of Illinois DCFS May Leave Amid Ethics Probe*, *chicagotribune.com*. Available at: https://www.chicagotribune.com/investigations/ct-dcfs-george-sheldon-met-20170503-story.html (Accessed: 9 September 2021).

Kelly, J. (2017) *Illinois Drops Rapid Safety Feedback, A Predictive Analytics Tool*, *The Imprint*. Available at: https://imprintnews.org/politics/stateline-illinois-drops-rapid-safety-feedback-predictive-analytics-tool/28913 (Accessed: 9 September 2021).

Office of Executive Inspector General (2017) 'Illinois Ethics Matters'.

The Executive Ethics Commission of the State of Illinois (2017) 'OEIG Final Report (Redacted)'. The Executive Ethics Commission of the State of Illinois.

New Zealand Vulnerable Children PRM (Predictive Risk Modelling)

Summary

The New Zealand government began exploring the potential of using a predictive risk modelling tool to identify children with the highest risk of neglect and abuse. As noted by Gillingham (Gillingham, 2019), the system was developed to combine multiple datasets and to identify the children in families with parents claiming public benefits who were most at risk of abuse and neglect. The stated aim was to provide supportive services to families.

Why was it stopped?

The system was trialled but never implemented. There was a great deal of critique as researchers and academics reviewed the system. Critique focused on the potential the system presented to further embed bias and lead to greater inequality in addition to concerns raised about accuracy and privacy. It was cancelled by Minister of Social Development Anne Tolley before an observational study was due to take place.

Key factors leading to change

The model was made public, and this provided an opportunity for academics and others to review it and raise concerns. Many were critical of the model. There was also critical media attention. A new Minister for Social Development, Anne Tolley, replaced the former minister and she has been credited as making the decision to shelve the system.

Ballantyne, who has researched this process, has argued that: "It is interesting to note that in spite of the ethical review commissioned by the MSD (Dare 2013) and the heated debates in academic journals (Dare, Vaithianathan & de Haan, 2014; Keddell, 2014a; Oak, 2015; Gillingham, 2015; Wilson et al. 2014; de Hann & Connolly, 2014) and news media, in the end New Zealand's experiment with predictive risk modelling in child protection services was closed down as the result of an intervention by a government Minister, it was a political decision. It is not possible to be certain of the detailed rationale for that decision, but the reasons the Minister gave to the press (Kirk, 2016), suggest that running an algorithm on all newborn children and intervening in cases not already known to social services – over half of which would be false positives – may have been a step too far for a neoliberal democracy" (Ballantyne 2019, p. 11).

How did the system work?

The system was developed by a team at Auckland University and relied on a statistical method to risk score children under the age of two for the likelihood they will experience harm or neglect within a population (Ballantyne interview, 2020). The idea was that those with high risk-scores would be targeted for in-home interventions to prevent maltreatment (Vaithianathan 2012; Vaithianathan et al. 2013; Wilson et al. 2015). The system based these scores on family histories and circumstances.

The system was announced as part of a 2012 government White Paper on Vulnerable Children and was the first time a government had announced an attempt to use predictive risk modelling for child maltreatment (Ballantyne interview, 2020). The model was based on New Zealand's social investment approach to welfare services which prioritized targeted practices instead of universal provision (Ballantyne interview, 2020).

The system made it through a series of trials, but the system was shelved before an observational study

was conducted by the new Minister for Social Development, Anne Tolley in 2015.

Impact

An ethical review was conducted which argued that the application of predictive risk modelling raised significant ethical concerns but that many of these concerns could be mitigated (Dare 2013). A feasibility study was also conducted.

In contrast, academics raised a number of concerns about the system targeting those who were poor, about accuracy and bias. Some of the academics to publicly criticize the system included: Patrick Kelly, Philip Gillingham, Emily Keddell, Eileen Oak and Ian Hyslop. The Social Workers Association and the Green Party were also critical.

There were criticisms that the system: presented too many false positives and that the data used in the system was inaccurate. There were criticisms that the system punished the poor as the highest weighted variables used were proxies for poverty which meant that low-income families would be disproportionately affected. Concerns were raised that this created a feedback loop as increasing surveillance would be focused on poor families. Questions were raised about what rights and protections families would have given the high number of inaccurate predictions (Gillingham 2015). Some questioned how many families would be worse off because of unjustified suspicion and the negative potential impact of stigmatizations (Oak 2016). Some raised concerns about how the system individualized social problems instead of addressing the structural causes of social problems (Keddell 2014).

Critiques of potential negative impacts were also publicized by mainstream media including Radio New Zealand and the New Zealand Herald.

Going forward

Some of the people involved in the development of this system went on to develop and implement predictive risk assessment tools that have been introduced elsewhere, such as in Alleghany, U.S.

Sources

Interviews conducted online 2020

Neil Ballantyne (Open Polytechnic of New Zealand)

Andrew Chen (University of Auckland)

Media: RNZ (Cowie, 2015)

Research: Keddell (2014) (Keddell, 2015), Gillingham (2019) (Gillingham, 2019), Gillingham (2016) (Gillingham, 2016)

Government: Ministry of Social Development (Ministry of Social Development, no date), Privacy Commissioner (Privacy Commissioner, New Zealand, 2015)

References

Ballantyne, N. (2019) 'The Ethics and Politics of Human Service Technology: The Case of Predictive Risk Modeling in New Zealand's Child Protection System', *The Hong Kong Journal of Social Work*, 53(01n02), pp. 15–27. doi:10.1142/S0219246219000044.

Bennett, P. (2012) *White Paper for Vulnerable Children, Launch, 11 October 2012*, *Beehive.gov.nz*. Available at: http://www.beehive.govt.nz/speech/white-paper-vulnerable-children-launch-11-october-2012 (Accessed: 13 September 2021).

Child Poverty Action Group (2017) *Proposed Funding Model Could Do More Harm*, *Child Poverty Action Group*. Available at: https://www.cpag.org.nz/news/proposed-funding-model-could-do-more-harm/ (Accessed: 13 September 2021).

Collins, S. (2012) *Predicting Trouble: Child Abuse Database Raises Eyebrows*, *NZ Herald*. Available at: https://www.nzherald.co.nz/nz/predicting-trouble-child-abuse-database-raises-eyebrows/UXLTJHIEKWAF5JDSZA7MPAHKQM/ (Accessed: 13 September 2021).

Cooke, H. (2019) *School Decile Funding System will be Scrapped by 2021 or 2022*, *Stuff*. Available at: https://www.stuff.co.nz/national/politics/116049946/school-decile-funding-system-will-be-scrapped-by-2021-or-2022 (Accessed: 13 September 2021).

Cowie, T. (2015) *Insight for 21 June 2015 - Child Abuse or Big Brother?*, *RNZ*. Available at: https://www.rnz.co.nz/national/programmes/insight/audio/201758628/insight-for-21-june-2015-child-abuse-or-big-brother (Accessed: 13 August 2021).

Dare, T. (2013) *Predictive Risk Modelling and Child Maltreatment: An Ethical Review*. University of Auckland.

Dare, T., Vaithianathan, R. and De Haan, I. (2014) Addressing Child Maltreatment in New Zealand: Is Poverty Reduction Enough?, *Educational Philosophy and Theory*, 46:9, 989-994, DOI: 10.1080/00131857.2014.938450

Gillingham, P. (2016) 'Predictive Risk Modelling to Prevent Child Maltreatment and Other Adverse Outcomes for Service Users: Inside the "Black Box" of Machine Learning', *The British Journal of Social Work*, 46(4), pp. 1044–1058. doi:10.1093/bjsw/bcv031.

Gillingham, P. (2019a) 'Decision Support Systems, Social Justice and Algorithmic Accountability in Social Work: A New Challenge', *Practice*, 31, pp. 1–14. doi:10.1080/09503153.2019.1575954.

Gillingham, P. (2019b) 'Decision Support Systems, Social Justice and Algorithmic Accountability in Social Work: A New Challenge', *Practice*, 31(4), pp. 277–290. doi:10.1080/09503153.2019.1575954.

Hyslop, I. (2013) 'The White Paper for Vulnerable Children and the Munro Review of Child Protection in England: a comparative critique', *Aotearoa New Zealand Social Work*, 25(4), pp. 4–14. doi:10.11157/anzswj-vol25iss4id58.

Jones, N. (2015) *Anne Tolley Scraps 'Lab Rat' Study on Children*, *NZ Herald*. Available at: https://www.nzherald.co.nz/nz/anne-tolley-scraps-lab-rat-study-on-children/C7GIGYW2467HG327FKXFRJDPEM/ (Accessed: 13 September 2021).

Keddell, E. (2015) 'The Ethics of Predictive Risk Modelling in the Aotearoa/New Zealand Child Welfare Context: Child Abuse Prevention or Neo-Liberal Tool?', *Critical Social Policy*, 35(1), pp. 69–88. doi:10.1177/0261018314543224.

Kilburn, M.R. and Karoly, L.A. (2008) *The Economics of Early Childhood Policy: What the Dismal Science Has to Say About Investing in Children*. RAND Corporation. Available at: https://www.rand.org/pubs/occasional_papers/OP227.html (Accessed: 13 September 2021).

Kirk, S. (2015) *Children 'Not Lab-Rats' - Anne Tolley Intervenes in Child Abuse Experiment*, *Stuff*. Available at: https://www.stuff.co.nz/national/health/70647353/children-not-lab-rats---anne-tolley-intervenes-in-child-abuse-experiment (Accessed: 13 September 2021).

Ministry of Social Development (no date) *Vulnerable Children Predictive Modelling*, *Ministry of Social Development, New Zealand*. Available at: https://www.msd.govt.nz/about-msd-and-our-work/publications-resources/research/predicitve-modelling/index.html (Accessed: 13 August 2021).

Ministry of Social Development, New Zealand (2014) *The Feasibility of Using Predictive Risk Modelling to Identify New-Born Children Who Are High Priority for Preventive Services*. Ministry of Social Development, New Zealand.

Oak, E. (2016) 'A Minority Report for Social Work? The Predictive Risk Model (PRM) and the Tuituia Assessment Framework in addressing the needs of New Zealand's Vulnerable Children', *The British Journal of Social Work*, 46(5), pp. 1208–1223. doi:10.1093/bjsw/bcv028.

Privacy Commissioner, New Zealand (2015) *Predicting Risk Without Turning Children Into 'Lab Rats'*, *Privacy Commissioner*. Available at: https://www.privacy.org.nz/blog/predictive-risk-modelling/ (Accessed: 13 August 2021).

Ryan, K. (2018) *Predictive Modelling: Digital vs Human Decision-Making*, *RNZ*. Available at: https://www.rnz.co.nz/national/programmes/ninetonoon/audio/2018664330/predictive-modelling-digital-vs-human-decision-making (Accessed: 13 September 2021).

Stuff.co.nz (2017) *Decision Close on Controversial Predictive Risk Modelling to Replace Decile Funding for Schools*, *Stuff*. Available at: https://www.stuff.co.nz/national/education/91476784/decision-close-on-controversial-predictive-risk-modelling-to-replace-decile-funding-for-schools (Accessed: 13 September 2021).

University of Auckland (2012) *Vulnerable Children: Can Administrative Data be Used to Identify Children at Risk of Adverse Outcomes?* Centre for Applied Research in Economics (CARE), University of Auckland.

Vaithianathan, R., Maloney, T., De Haan, I., & Dare, T. (2012). Vulnerable children: Can administrative data be used to identify children at risk of adverse outcomes? Auckland. Retrieved from http://www.msd.govt.nz/about-msd-and-our-work/publications-resources/research/vulnerable-children/index.html

Vaithianathan, R., Maloney, T., Putnam-Hornstein, E., & Jiang, N. (2013). Children in the public benefit system at risk of maltreatment: Identification via predictive modeling. American Journal of Preventive Medicine, 45(3), 354–359. Retrieved from http://www.sciencedirect.com/science/article/pii/S0749379713003449

Wilson, M. L., Tumen, S., Ota, R., & Simmers, A. G. (2015). Predictive modeling: Potential application in prevention services. American Journal of Preventive Medicine, 48(5), 509–519. https://doi.org/10.1016/j.amepre.2014.12.003

Hackney Council stops Early Help Profiling System

Summary

In 2015, Hackney Council, UK, started using Xantura's Early Help Profiling System (EHPS) to risk assess at-risk families. The predictive system would send an alert and a report to case workers if a risk threshold had been crossed. The pilot scheme was dropped in 2019.

Why was it stopped?

A Hackney Council spokesperson is quoted in the Hackney Citizen as saying: "At the conclusion of the pilot we had not been able to realise the expected benefits and decided to not continue beyond the pilot stage. We found that the data available was more limited than had initially been envisaged and issues of variable data quality meant that the system wasn't able to provide sufficiently useful insights to justify further investment in the project." Concerns were raised about privacy and consent in media coverage as well as by a local politician.

Key factors leading to change

It is difficult to determine as there is very little information. Reporting suggests that it relates to effectiveness and data quality and that it was an internal decision. Specifically, it was reported that EHPS wasn't compatible with other council systems, that data collection was difficult, that expected benefits were not realized and that data quality was variable (Sheridan 2019).

How did it work?

The EHPS model analyses multiple data sources, including school attendance and attainment, health records, families' housing situations, and economic indicators, to assess families' risk scores. The EHPS identifies and alerts social workers to those families who are flagged as needing extra support. Each month the system provides the council's social workers with a list of 20 families whose risk score indicates they are most in need. It has been emphasized in publications about the system that it has been designed not to replace the decision making of social workers but to support them (Apolitical, 2017).

London Councils, the umbrella body that oversees London's 32 councils, was the actor responsible for bringing in the trial of EHPS. All of the councils in London were invited to participate. Hackney, Thurrock, Newham and Tower Hamlets participated (Apolitical, 2017). EHPS was promoted to help Councils deliver the Troubled Families program, which had funding issues as reported by the Hackney Citizen in Feb 2019 (Sheridan, 2019). After two years of testing EHPS Hackney Council began using it monthly to generate a list of between 10 and 20 families flagged for future concern (Stevenson, 2018). As part of the project, Hackney Council was due to run a pilot to share the EHPS data with GPs in Autumn 2017 "to assist them with making referrals to children's social care, which they will be able to make through this system directly to our front door." (Apolitical, 2017; Community Care, 2018)

Development and testing of EHPS was funded by the London Ventures innovation programme, a partnership between umbrella body London Councils and professional services firm EY (Apolitical 2017; Children and Young People Now, Stephenson, 2018). One of the rationales for introducing the system was to save money through early targeted interventions, replacing human-conducted screenings with an automated system and improving access to data. The need to save money was driven, in part, by an austerity context that saw councils in England have their funding from central government cut by nearly a third in real terms since the Conservatives launched their austerity program in 2010 (Innes and Tetlow 2015).

Impact

It was reported in 2018 that EHPS alerts had led to early interventions and support for 400 families through the Troubled Families programme. It was also reported that the system reduced the time needed to process and research information (Community Care 2018). Other research noted a lack of information available about impact on affected people and decision makers (Dencik et al. 2018).

Privacy concerns were raised by the Liberal Democrats about the decision to not share with citizens that their data was being used in the system. Xantura, the company that developed the system, argued it would reduce functionality if people knew their data was being used. It was also argued that specific details about how the system worked could not be shared due to commercial sensitivity.

Going forward

London Councils has been promoting Xantura's predictive technologies as a way for local authorities to address needs during Covid-19

References

Apolitical (2017) *London Uses Data to Predict Which Children will be Abused*, *Apolitical*. Available at: https://apolitical.co/solution-articles/en/london-uses-data-predict-which-children-abuse (Accessed: 13 September 2021).

Hackney Borough Council (2018) *Documents Relating to the Children's Safeguarding Profiling System - a Freedom of Information Request to Hackney Borough Council*, *WhatDoTheyKnow*. Available at: https://www.whatdotheyknow.com/request/documents_relating_to_the_childr (Accessed: 13 September 2021).

Innes, David, & Tetlow, Gemma. (2015, September). Delivering fiscal squeeze by cutting local government spending. Fiscal Studies, 36(3), 303–325. URL: https://www.ifs.org.uk/publications /8033 [October 5, 2018].

London Councils (no date a) *Understanding Who the Most Vulnerable People in Your Locality Are*. Available at: https://www.londoncouncils.gov.uk/our-key-themes/our-projects/london-ventures/current-projects/childrens-safeguarding (Accessed: 13 September 2021).

London Councils (no date b) *Venture Spotlight: Xantura*. Available at: https://www.londoncouncils.gov.uk/node/31412 (Accessed: 13 September 2021).

Say, M. (2018) *Hackney Council Automates FoI Responses*, *UKAuthority*. Available at: https://www.ukauthority.com/articles/hackney-council-automates-foi-responses/ (Accessed: 13 September 2021).

Sheridan, E. (2018) 'Hackney Council Pays £360k to Data Firm Whose Software Profiles Troubled Families', *Hackney Citizen*, 18 October. Available at: https://www.hackneycitizen.co.uk/2018/10/18/council-360k-xantura-software-profiles-troubled-families/ (Accessed: 13 September 2021).

Sheridan, E. (2019a) 'Council's Troubled Families Programme Under Threat Amid Funding Uncertainty', *Hackney Citizen*, 21 February. Available at: https://www.hackneycitizen.co.uk/2019/02/21/councils-troubled-families-programme-under-threat-funding-uncertainty/ (Accessed: 13 September 2021).

Sheridan, E. (2019b) 'Town Hall Drops Pilot Programme Profiling Families Without Their Knowledge', *Hackney Citizen*, 30 October. Available at: https://www.hackneycitizen.co.uk/2019/10/30/town-hall-drops-pilot-programme-profiling-families-without-their-knowledge/ (Accessed: 15 July 2021).

Stephenson, J. (2018) *Early Help Profiling System*, *CYP Now*. Available at: https://www.cypnow.co.uk/best-practice/article/early-help-profiling-system (Accessed: 13 September 2021).

Stevenson, L. (2018) 'Artificial Intelligence: How a Council Seeks to Predict Support Needs for Children and Families', *Community Care*, 1 March. Available at: https://www.communitycare.co.uk/2018/03/01/artificial-intelligence-council-seeks-predict-support-needs-children-families/ (Accessed: 13 September 2021).

Bibliography

Adams, R., Weale, S. and Barr, C. (2020) *A-Level Results: Almost 40% of Teacher Assessments in England Downgraded, the Guardian*. Available at: http://www.theguardian.com/education/2020/aug/13/almost-40-of-english-students-have-a-level-results-downgraded (Accessed: 15 July 2021).

AI Now Institute Report (2018) *Litigating Algorithms: Challenging Government Use of Algorithmic Decision Systems*, p. 17. Available at: https://ainowinstitute.org/litigatingalgorithms.pdf.

Alexander, R. (2017) *Spokane Judges Have a New Tool to Decide Whether the Right People are in Jail, The Spokesman-Review*. Available at: https://www.spokesman.com/stories/2017/jan/11/spokane-judges-have-a-new-tool-to-decide-whether-t/ (Accessed: 7 July 2021).

Algorithm Watch (2019) 'Automating Society', *AlgorithmWatch*. Available at: https://algorithmwatch.org/en/automating-society/ (Accessed: 18 August 2021).

Allen, K. (2020) 'Toronto Police Chief Halts Use of Controversial Facial Recognition Tool', *thestar.com*, 13 February. Available at: https://www.thestar.com/news/gta/2020/02/13/toronto-police-used-clearview-ai-an-incredibly-controversial-facial-recognition-tool.html (Accessed: 8 July 2021).

Allen, K. and Gillis, W. (2019) 'Toronto Police have been Using Facial Recognition Technology for more than a Year', *thestar.com*, 28 May. Available at: https://www.thestar.com/news/gta/2019/05/28/toronto-police-chief-releases-report-on-use-of-facial-recognition-technology.html (Accessed: 8 July 2021).

Allen, K., Gillis, W. and Boutilier, A. (2020) *Facial Recognition App Clearview AI Has Been Used Far More Widely In Canada Than Previously Known, thestar.com*. Available at: https://www.thestar.com/news/canada/2020/02/27/facial-recognition-app-clearview-ai-has-been-used-far-more-widely-in-canada-than-previously-known.html (Accessed: 8 July 2021).

Alston, P. (2018) *Statement on Visit to the United Kingdom, by Professor Philip Alston, United Nations Special Rapporteur on extreme poverty and human rights, United Nations Human Rights Office of the High Commissioner*. Available at: https://www.ohchr.org/en/NewsEvents/Pages/DisplayNews.aspx?NewsID=23881 (Accessed: 18 August 2021).

Altermatt, S. (2017) *Kein klarer Nutzen: Solothurner Polizei verzichtet auf Einbruchssoftware, Solothurner Zeitung*. Available at: https://www.solothurnerzeitung.ch/solothurn/kanton-solothurn/kein-klarer-nutzen-solothurner-polizei-verzichtet-auf-einbruchssoftware-ld.1471709 (Accessed: 15 July 2021).

American Federation of Teachers (2017) *Federal Suit Settlement: End of Value-Added Measures in Houston.* Available at: https://www.educationviews.org/federal-suit-settlement-value-added-measures-houston/, https://www.educationviews.org/federal-suit-settlement-value-added-measures-houston/ (Accessed: 29 June 2021).

Amrein-Beardsley, A. (2019) 'The Education Value-Added Assessment System (EVAAS) on Trial: A Precedent-Setting Lawsuit with Implications for Policy and Practice', *eJournal of Education Policy*.

van Ark, T. (2020) 'Letter of 23 April 2020 to the President of the House of Representatives by the State Secretary for Social Affairs and Employment, Tamara van Ark, on a Court Judgment Regarding SyRI'.

Bailey, L. (2017) 'Software Firms Sued Over Mich. Jobless System', *Courthouse News Service*, 3 March. Available at: https://www.courthousenews.com/software-firms-sued-michigan-jobless-system/ (Accessed: 29 June 2021).

Barbaschow, A. (2019) *Government Backflip As Robo-Debt Income Automation Paused*, *ZDNet*. Available at: https://www.zdnet.com/article/government-backflip-as-robo-debt-income-automation-paused/ (Accessed: 12 August 2021).

Barrett, L. (2020) 'Ban Facial Recognition Technologies for Children – And for Everyone Else', *BUJ Sci. & Tech. L.* 26: 223-285.

Barrett, M. (2017) *After Years of Scandals, Here's How Denmark's Reformed Tax Authority Will Look*, *The Local*. Available at: https://www.thelocal.dk/20170614/heres-how-denmarks-reformed-tax-administration-will-look/ (Accessed: 14 July 2021).

BBC News (2018) 'Kent Police Stop Using Crime Predicting Software', *BBC News*, 26 November. Available at: https://www.bbc.com/news/uk-england-kent-46345717 (Accessed: 15 July 2021).

BBC News (2020) 'Home Office Drops "Racist" Algorithm from Visa Decisions', *BBC News*, 4 August. Available at: https://www.bbc.com/news/technology-53650758 (Accessed: 11 August 2021).

Bertsimas, D. *et al.* (2020) 'Bus Routing Optimization Helps Boston Public Schools Design Better Policies', *INFORMS Journal on Applied Analytics*, 50(1), pp. 37–49. doi: 10.1287/inte.2019.1015.

Big Brother Watch (2018) *A Closer Look at Experian Big Data and Artificial Intelligence in Durham Police*, *Big Brother Watch*. Available at: https://bigbrotherwatch.org.uk/2018/04/a-closer-look-at-experian-big-data-and-artificial-intelligence-in-durham-police/#_ftn2 (Accessed: 15 July 2021).

Bij Voorbaat Verdacht (2019) 'Save the Date: Datum zitting bodemprocedure tegen SyRI op 29 oktober 2019', *Bij Voorbaat Verdacht*, 23 April. Available at: https://bijvoorbaatverdacht.nl/save-the-date-zitting-rechtszaak-tegen-syri-op-29-oktober-2019/ (Accessed: 14 July 2021).

Bonnett, G. (2018a) *Immigration Dumps Controversial Deportation Analytical Tool*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/361199/immigration-dumps-controversial-deportation-analytical-tool (Accessed: 12 August 2021).

Bonnett, G. (2018b) *Immigration NZ Using Data System to Predict Likely Troublemakers*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/354135/immigration-nz-using-data-system-to-predict-likely-troublemakers (Accessed: 12 August 2021).

Bracewell-Worrall, A. (2018) 'Immigration Minister Puts "Ethnic Profiling" Pilot Programme on Hold', *Newshub*, 4 September. Available at: https://www.newshub.co.nz/home/politics/2018/04/immigration-minister-puts-ethic-profiling-pilot-programme-on-hold.html (Accessed: 12 August 2021).

Brandimarte, E. (2019) *Should We Let the Algorithms Decide? A Critical Assessment of Automated Decision-Making in Europe*. College of Europe.

Bristol Government (2020) *Review and Rationale for Replacement of RBV*.

Brusselsexpress.eu (2019) *Facial Recognition Cameras at Brussels Airport "Have No Legal Basis", Says Lawyer*, *Brussels Express*. Available at: https://brussels-express.eu/facial-recognition-cameras-at-brussels-airport-have-no-legal-basis-says-lawyer/ (Accessed: 8 July 2021).

Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77-91). *PMLR*.

Burgess, M. (2018) *UK Police are Using AI to Make Custodial Decisions – But It Could Be Discriminating Against the Poor*, *WIRED UK*. Available at: https://www.wired.co.uk/article/police-ai-uk-durham-hart-checkpoint-algorithm-edit (Accessed: 15 July 2021).

Burgess, M. (2020) 'Police built an AI to predict violent crime. It was seriously flawed', *Wired UK*, 8 June. Available at: https://www.wired.co.uk/article/police-violence-prediction-ndas (Accessed: 12 August 2021).

'California Governor Signs Landmark Bill Halting Facial Recognition on Police Body Cams' (2019). ACLU of Northern CA. Available at: https://www.aclunc.org/news/california-governor-signs-landmark-bill-halting-facial-recognition-police-body-cams (Accessed: 7 July 2021).

Callahan, R. (2018) *Appeals Court Affirms Ruling That IBM Owes Indiana $78M*, *AP News*. Available at: https://apnews.com/article/36ba0562a02142e5adfe39518e2e0f85 (Accessed: 4 August 2021).

Cardozo, T. and Freeze, C. (2020) 'Toronto Police Chief Orders Officers to Stop using Clearview AI Software', 13 February. Available at: https://www.theglobeandmail.com/canada/toronto/article-toronto-police-chief-orders-officers-to-stop-using-clearview-ai/ (Accessed: 8 July 2021).

Carney, B. (2020) *FOI Documents Confirm RCMP Falsely Denied Using Facial Recognition Software*, *The Tyee*. The Tyee. Available at: https://thetyee.ca/News/2020/04/16/RCMP-FOI-Facial-Recognition/ (Accessed: 8 July 2021).

Carney, T. (2019) 'Robo-Debt Illegality: The Seven Veils of Failed Guarantees of the Rule of Law?', *Alternative Law Journal*, 44(1), pp. 4–10. doi: 10.1177/1037969X18815913.

CBC News (2012) *Police Can't Use ICBC Facial Recognition to Track Rioters*, *CBC*. Available at: https://www.cbc.ca/news/canada/british-columbia/police-can-t-use-icbc-facial-recognition-to-track-rioters-1.1207398 (Accessed: 8 July 2021).

CBC News (2020) *Toronto Police Admit Using Secretive Facial Recognition Technology Clearview AI*, *CBC*. Available at: https://www.cbc.ca/news/canada/toronto/toronto-police-clearview-ai-1.5462785 (Accessed: 8 July 2021).

Chan, J. B. (2016) *ShotSpotter Not Exactly Taking a Bite Out Of Crime*, *Reveal*. Available at: http://revealnews.org/blog/shotspotter-not-exactly-taking-a-bite-out-of-crime/ (Accessed: 8 July 2021).

Charette, R. N. (2012) *Troubled HealthSMART System Finally Cancelled in Victoria Australia*, *IEEE Spectrum*. Available at: https://spectrum.ieee.org/troubled-healthsmart-system-finally-cancelled-in-victoria-australia- (Accessed: 12 August 2021).

Charette, R. N. (2018a) *Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold*, *IEEE Spectrum*. Available at: https://spectrum.ieee.org/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold (Accessed: 4 August 2021).

Charette, R. N. (2018b) *Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold - IEEE Spectrum*, *IEEE Spectrum: Technology, Engineering, and Science News*. Available at: https://spectrum.ieee.org/riskfactor/computing/software/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold (Accessed: 29 June 2021).

Chiusi, F. (2020) *Automating Society Report 2020*. Berlin, Germany: AlgorithmWatch, p. 298. Available at: https://automatingsociety.algorithmwatch.org (Accessed: 14 July 2021).

Chowdhury, H. (2018) *Kent Police Stop Using Crime Predicting Software*, *The Telegraph*. Available at: https://www.telegraph.co.uk/technology/2018/11/27/kent-police-stop-using-crime-predicting-software/ (Accessed: 15 July 2021).

Clarke, T. (2010) *Victorian Police Suspend Crime Database Overhaul*, *Computerworld*. Available at: https://www.computerworld.com/article/3462991/victorian-police-suspend-crime-database-overhaul.html (Accessed: 12 August 2021).

'CNIL Bans High Schools' Facial-Recognition Programs' (2019) *iapp*, 29 October. Available at: https://iapp.org/news/a/cnil-bans-high-school-facial-recognition-programs/ (Accessed: 14 July 2021).

Coleman, E. (2020) *One City Rejected a Policing Algorithm. Could It Be Used For Other Purposes?*, *Route Fifty*. Available at: https://www.route-fifty.com/public-safety/2020/07/pittsburgh-hot-spot-algorithm/166953/ (Accessed: 8 July 2021).

Commonwealth Ombudsman (2019) *Submission by the Commonwealth Ombudsman: Inquiry Into Centrelink's Compliance Program*.

Conger, K., Fausset, R. and Kovaleski, S. F. (2019) 'San Francisco Bans Facial Recognition Technology', *The New York Times*, 14 May. Available at: https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html (Accessed: 7 July 2021).

Coop, A. (2020) *Privacy Expert Calls Toronto Police's Use of Clearview AI 'Appalling'*, *IT World Canada News*. Available at: https://www.itworldcanada.com/article/privacy-expert-calls-toronto-polices-use-of-clearview-ai-appalling/427346 (Accessed: 8 July 2021).

Coughlan, S. (2020) 'Why did the A-level algorithm say no?', *BBC News*, 14 August. Available at: https://www.bbc.com/news/education-53787203 (Accessed: 15 July 2021).

Counsel to the Attorney General (2020) 'Cease and Desist'.

County of Los Angeles Office of Child Protection (2016) *Examination of Using Structured Decision Making and Predictive Analytics in Assessing Safety and Risk in Child Welfare*. County of Los Angeles, Office of Child Protection, p. 18.

Cournoyer, C. (2017) *Too Much Data? Illinois Abandons System Meant to Predict Child Abuse*, *Governing: The Future of States and Localities*. Available at: https://www.governing.com/archive/tns-chicago-data-mining.html (Accessed: 29 June 2021).

Cowan, P. (2015) *AFP Dumps $145m Investigation System Replacement*, *iTnews*. Available at: https://www.itnews.com.au/news/afp-dumps-145m-investigation-system-replacement-408034 (Accessed: 12 August 2021).

Cowie, T. (2015) *Insight for 21 June 2015 - Child Abuse or Big Brother?, RNZ*. Available at: https://www.rnz.co.nz/national/programmes/insight/audio/201758628/insight-for-21-june-2015-child-abuse-or-big-brother (Accessed: 13 August 2021).

Crenshaw, K. (2015) 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', *University of Chicago Legal Forum*, 1989(1). Available at: https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8.

Crist, B. (2019) 'An Important Update from Chef', *Chef Blog*, 23 September. Available at: https://blog.chef.io/an-important-update-from-chef (Accessed: 7 July 2021).

Crozier, R. (2012) *Victoria kills HealthSMART IT project*, *iTnews*. Available at: https://www.itnews.com.au/news/victoria-kills-healthsmart-it-project-301410 (Accessed: 12 August 2021).

Data Justice Lab (2018) *Towards Democratic Auditing: Civic Participation in the Scoring Society*, *Data Justice Lab*. Available at: https://datajusticelab.org/towards-democratic-auditing-civic-participation-in-the-scoring-society/ (Accessed: 11 August 2021).

'David Mutnick, for himself and others similarly situated, vs. Clearview AI, Inc., Hoan Ton-That and Richard Schwartz' (2020). US District Court for the Northern District of Illinois Eastern Division. Available at: https://www.scribd.com/document/444154093/gov-uscourts-ilnd-372790-1-0 (Accessed: 8 July 2021).

de Rechtspraak (2020) *SyRI-wetgeving in strijd met het Europees Verdrag voor de Rechten voor de Mens*, *de Rechtspraak*. Available at: https://www.rechtspraak.nl/Organisatie-en-contact/Organisatie/Rechtbanken/Rechtbank-Den-Haag/Nieuws/Paginas/SyRI-wetgeving-in-strijd-met-het-Europees-Verdrag-voor-de-Rechten-voor-de-Mens.aspx (Accessed: 14 July 2021).

Department for Work and Pensions (2011) *Housing Benefit and Council Tax Benefit Circular*. HB/CTB S11/2011. London.

Devanesan, J. (2020) *New Zealand Police Trial Controversial Clearview AI Facial Recognition Tech*, *Tech Wire Asia*. Available at: https://techwireasia.com/2020/05/police-trial-controversial-clearview-ai-facial-recognition-tech-in-new-zealand/ (Accessed: 13 August 2021).

Digital Society Blog (2018) 'Profiling the Unemployed', *HIIG*, 16 January. Available at: https://www.hiig.de/en/profiling-the-unemployed/ (Accessed: 15 July 2021).

Eaton, M. and Bertoncin, C. (2018) *State of Offices of Data Analytics (ODA) in the UK*. Nesta.

Edes, A. and Bowman, E. (2018) *'Automating Inequality': Algorithms In Public Services Often Fail The Most Vulnerable*, *NPR.org*. Available at: https://www.npr.org/sections/alltechconsidered/2018/02/19/586387119/automating-inequality-algorithms-in-public-services-often-fail-the-most-vulnerab (Accessed: 29 June 2021).

Egbert, S. and Krasmann, S. (2019) *Predictive Policing. Eine ethnographische Studie neuer Technologien zur Vorhersage von Straftaten und ihre Folgen für die polizeiliche Praxis*. Hamburg: Universität Hamburg.

Eubanks, V. (2015) 'Want to Cut Welfare? There's an App for That', 27 May. Available at: https://www.thenation.com/article/archive/want-cut-welfare-theres-app/ (Accessed: 29 June 2021).

Eubanks, V. (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. First Edition. New York: St Martin's Press.

European Court of Human Rights (2020) *European Convention on Human Rights*. France: Council of Europe, p. 34.

Final Report of the Special Commission to Evaluate Policies and Procedures Related to the Current Bail System (2019). Government document. Boston: The General Court of Massachusetts State House, p. 114.

Finley, K. (2019) 'Software Company Chef Won't Renew ICE Contact After All', *Wired*, 23 September. Available at: https://www.wired.com/story/software-company-chef-wont-renew-ice-contact/ (Accessed: 7 July 2021).

Foxglove (2020) *Home Office Says it Will Abandon its Racist Visa Algorithm - After We Sued Them*, *Foxglove*. Available at: https://www.foxglove.org.uk/2020/08/04/home-office-says-it-will-abandon-its-racist-visa-algorithm-after-we-sued-them/ (Accessed: 11 August 2021).

Galindo, G. (2019) *'No legal basis' for facial recognition cameras at Brussels Airport*, *The Brussels Times*. Available at: https://www.brusselstimes.com/brussels-2/60362/no-legal-basis-for-facial-recognition-cameras-identity-brussels-airport-intelligent-cameras/ (Accessed: 8 July 2021).

Gavaghan, C. (2018) *Where Did it Algo wrong? The Threat and Promise of Predictive Analytics*, *Pundit*. Available at: https://www.pundit.co.nz/content/where-did-it-algo-wrong-the-threat-and-promise-of-predictive-analytics (Accessed: 12 August 2021).

Gerstner, D. (2017) *Predictive Policing als Instrument zur Prävention von Wohnungseinbruchdiebstahl: Evaluationsergebnisse zum baden-württembergischen Pilotprojekt P4*. Freiburg im Breisgau: Max-Planck-Institut für ausländisches und internationales Strafrecht.

Gerstner, D. (2018) 'Predictive Policing in the Context of Residential Burglary: An Empirical Illustration on the Basis of a Pilot Project in Baden-Württemberg, Germany', *European Journal for Security Research*, 3(2), pp. 115–138. doi: 10.1007/s41125-018-0033-0.

Ghaffary, S. (2019) *San Francisco's Facial Recognition Technology Ban, Explained*, *Vox*. Available at: https://www.vox.com/recode/2019/5/14/18623897/san-francisco-facial-recognition-ban-explained (Accessed: 7 July 2021).

Gillingham, P. (2016) 'Predictive Risk Modelling to Prevent Child Maltreatment and Other Adverse Outcomes for Service Users: Inside the "Black Box" of Machine Learning', *The British Journal of Social Work*, 46(4), pp. 1044–1058. doi: 10.1093/bjsw/bcv031.

Gillingham, P. (2019) 'Decision Support Systems, Social Justice and Algorithmic Accountability in Social Work: A New Challenge', *Practice*, 31(4), pp. 277–290. doi: 10.1080/09503153.2019.1575954.

Gonzalez, M. G. (2020) *Cambridge City Council Votes to Ban Facial Recognition Technology*, *The Harvard Crimson*. Available at: https://www.thecrimson.com/article/2020/1/16/cambridge-city-council-bans-facial-recognition/ (Accessed: 7 July 2021).

Gonzalez, S. (2013) *In Newark, Gunshot Detection System Falls Short of Booker's Claims*, *WNYC*. Available at: https://www.wnyc.org/story/311533-gunshot-detection-sensors-newark-result-17-arrests-over-three-

years/ (Accessed: 8 July 2021).

Government of Canada, O. of the A. G. of C. (2017) *Report 1—Phoenix Pay Problems*. Available at: https://www.oag-bvg.gc.ca/internet/English/parl_oag_201711_01_e_42666.html (Accessed: 8 July 2021).

Government of Canada, O. of the A. G. of C. (2018) *Report 1—Building and Implementing the Phoenix Pay System*. Available at: https://www.oag-bvg.gc.ca/internet/English/parl_oag_201805_01_e_43033.html (Accessed: 8 July 2021).

Gray, J. (2019) 'Toronto Police End ShotSpotter Project Over Legal Concerns', *The Globe and Mail*, 13 February. Available at: https://www.theglobeandmail.com/canada/toronto/article-toronto-police-end-shotspotter-project-over-legal-concerns/ (Accessed: 8 July 2021).

Guariglia, M. (2019) *Victory! Berkeley City Council Unanimously Votes to Ban Face Recognition*, *Electronic Frontier Foundation*. Available at: https://www.eff.org/deeplinks/2019/10/victory-berkeley-city-council-unanimously-votes-ban-face-recognition (Accessed: 7 July 2021).

Haag, D. (2020) *SyRI-wetgeving in strijd met het Europees Verdrag voor de Rechten voor de Mens*. Available at: https://www.rechtspraak.nl/Organisatie-en-contact/Organisatie/Rechtbanken/Rechtbank-Den-Haag/Nieuws/Paginas/SyRI-wetgeving-in-strijd-met-het-Europees-Verdrag-voor-de-Rechten-voor-de-Mens.aspx (Accessed: 18 August 2021).

Haskins, C. (2019) *Oakland Becomes Third U.S. City to Ban Facial Recognition*, *Vice*. Available at: https://www.vice.com/en/article/zmpaex/oakland-becomes-third-us-city-to-ban-facial-recognition-xz (Accessed: 7 July 2021).

Hegarty, P. (2019) *East Bay City Becomes Latest to Ban Use of Facial Recognition Technology*, *East Bay Times*. Available at: https://www.eastbaytimes.com/2019/12/18/east-bay-city-becomes-latest-to-ban-use-of-facial-recognition-technology (Accessed: 7 July 2021).

Hendry, J. (2018) *AFP Revives Investigation System Replacement*, *iTnews*. Available at: https://www.itnews.com.au/news/afp-revives-investigation-system-replacement-489890 (Accessed: 12 August 2021).

Henriques-Gomes, L. (2019) *Centrelink Still Issuing Incorrect Robodebts To Meet Targets, Staff Claim*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2019/may/29/centrelink-still-issuing-incorrect-robodebts-to-meet-targets-staff-claim (Accessed: 12 August 2021).

Henriques-Gomes, L. (2020) *Officials Refuse To Say How Many Flawed Robodebts They Have Found*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2020/mar/06/centrelink-robodebt-senate-estimates-class-action (Accessed: 12 August 2021).

Hern, A. (2020) *Ofqual's A-Level Algorithm: Why Did it Fail to Make the Grade?*, *the Guardian*. Available at: http://www.theguardian.com/education/2020/aug/21/ofqual-exams-algorithm-why-did-it-fail-make-grade-a-levels (Accessed: 15 July 2021).

Hesse, D. and Zürcher, C. (2018) 'Du wirst verbrochen haben', *Der Bund*, 11 May. Available at: https://www.derbund.ch/schweiz/standard/du-wirst-verbrochen-haben/story/13799968 (Accessed: 15 July 2021).

Heuberger, von S. (2017) *Predictive Policing: Software ist keine Wunderwaffe gegen Einbrüche*, *GQ*. Available at: https://www.gq-magazin.de/auto-technik/article/mit-big-data-den-einbrecher-schnappen-klappt-

nur-bedingt (Accessed: 14 July 2021).

Hill, K. (2020) 'New Jersey Bars Police From Using Clearview Facial Recognition App', *The New York Times*, 25 January. Available at: https://www.nytimes.com/2020/01/24/technology/clearview-ai-new-jersey.html (Accessed: 8 July 2021).

Hoffmann, A. L. (2019) 'Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse', *Information, Communication & Society*, 22(7), pp. 900–915. doi: 10.1080/1369118X.2019.1573912.

Houston Federation of Teachers, Local 2415, et al., versus Houston Independent School District (2017).

Howgego, J. (2019) *A UK Police Force is Dropping Tricky Cases on Advice of an Algorithm*, *New Scientist*. Available at: https://www.newscientist.com/article/2189986-a-uk-police-force-is-dropping-tricky-cases-on-advice-of-an-algorithm/ (Accessed: 15 July 2021).

Human Rights Ombudsman (2018) *Annual Report of the Human Rights Ombudsman of the Republic of Slovenia for 2017*. 23rd. Ljubljana, Slovenia, p. 256.

Ito, J. (2018) 'Appendix B: Letter 1 to Massachusetts Legislature on Pretrial Risk Assessment', *Practice of Change*. Available at: https://www.practiceofchange.org/pub/n2wzxaft/release/1 (Accessed: 8 July 2021).

Jackson, D. and Marx, G. (2017) *Data Mining Program Designed to Predict Child Abuse Proves Unreliable, DCFS says*, *Chicago Tribune*. Available at: https://www.chicagotribune.com/investigations/ct-dcfs-eckerd-met-20171206-story.html (Accessed: 29 June 2021).

Jarmanning, A. (2020) *Boston Lawmakers Vote To Ban Use Of Facial Recognition Technology By The City*, *NPR.org*. Available at: https://www.npr.org/sections/live-updates-protests-for-racial-justice/2020/06/24/883107627/boston-lawmakers-vote-to-ban-use-of-facial-recognition-technology-by-the-city (Accessed: 7 July 2021).

JCWI (no date) *We Won! Home Office to Stop Using Racist Visa Algorithm*, *The Joint Council for the Welfare of Immigrants*. Available at: https://www.jcwi.org.uk/news/we-won-home-office-to-stop-using-racist-visa-algorithm (Accessed: 11 August 2021).

Karp, P. (2019a) *Government Admits Robodebt Was Unlawful As It Settles Legal Challenge*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2019/nov/27/government-admits-robodebt-was-unlawful-as-it-settles-legal-challenge (Accessed: 12 August 2021).

Karp, P. (2019b) *Robodebt: Government Abandons Key Part Of Debt Recovery Scheme In Major Overhaul*, *the Guardian*. Available at: http://www.theguardian.com/australia-news/2019/nov/19/robodebt-government-abandons-key-part-of-debt-recovery-scheme-in-major-overhaul (Accessed: 12 August 2021).

Keddell, E. (2015) 'The Ethics of Predictive Risk Modelling in the Aotearoa/New Zealand Child Welfare Context: Child Abuse Prevention or Neo-Liberal Tool?', *Critical Social Policy*, 35(1), pp. 69–88. doi: 10.1177/0261018314543224.

Kjær, J. S. (2018) *Regeringen har lagt sin plan om overvågning af børnefamilier i skuffen - politiken.dk*, *Politiken*. Available at: https://politiken.dk/indland/art6919255/Regeringen-har-lagt-sin-plan-om-overv%C3%A5gning-af-b%C3%B8rnefamilier-i-skuffen (Accessed: 14 July 2021).

Kolkman, D. (2020) '"F**k the Algorithm"?: What the World can Learn from the UK's A-Level Grading Fiasco', *Impact of Social Sciences*, 26 August. Available at: https://blogs.lse.ac.uk/impactofsocialsciences/2020/08/26/fk-the-algorithm-what-the-world-can-learn-from-the-uks-a-level-grading-fiasco/ (Accessed: 15 July 2021).

Lee, D. (2019) 'San Francisco is First US City to Ban Facial Recognition', *BBC News*, 14 May. Available at: https://www.bbc.com/news/technology-48276660 (Accessed: 7 July 2021).

'Letter to Los Angeles Police Commission' (2019). Available at: https://stoplapdspying.medium.com/on-tuesday-april-2nd-2019-twenty-eight-professors-and-forty-graduate-students-of-university-of-8ed7da1a8655 (Accessed: 7 July 2021).

Lipton, B. (2019) *Five Cities, Five Different Approaches to Algorithms in Government*, *MuckRock*. Available at: https://www.muckrock.com/news/archives/2019/jul/23/algorithms-five-cities/ (Accessed: 7 July 2021).

Mac, R., Haskins, C. and McDonald, L. (2020) *Clearview's Facial Recognition App Has Been Used By The Justice Department, ICE, Macy's, Walmart, And The NBA*, *BuzzFeed News*. Available at: https://www.buzzfeednews.com/article/ryanmac/clearview-ai-fbi-ice-global-law-enforcement (Accessed: 8 July 2021).

Macias Jr., M. (2019) *LAPD Says It Will Dump Data-Driven Policing Tools*, *Courthouse News Service*. Available at: https://www.courthousenews.com/lapd-says-it-will-dump-data-driven-policing-tools/ (Accessed: 7 July 2021).

Mackenzie, S. (2020a) *Police 'Stocktake' Surveillance Tech After Clearview AI Facial Recognition Trial*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/416913/police-stocktake-surveillance-tech-after-clearview-ai-facial-recognition-trial (Accessed: 13 August 2021).

Mackenzie, S. (2020b) *Police Trialled Facial Recognition Tech Without Clearance*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/416483/police-trialled-facial-recognition-tech-without-clearance (Accessed: 13 August 2021).

Markey, E. J. (2020) 'Letter to Mr. Hoan Ton-That'. US Senate.

Marsh, S. (2019) *One in Three Councils Using Algorithms to Make Welfare Decisions*, *the Guardian*. Available at: http://www.theguardian.com/society/2019/oct/15/councils-using-algorithms-make-welfare-decisions-benefits (Accessed: 15 July 2021).

Mayer, R. W. N. (2017) *Einbrüche in Stuttgart: Software soll Einbrecher stoppen*, *stuttgarter-nachrichten.de*. Available at: https://www.stuttgarter-nachrichten.de/inhalt.land-testet-einbruchs-software-kommissar-computer.ee56faed-3723-405d-8986-9aa76dfc888d.html (Accessed: 14 July 2021).

McDonald, H. (2020) *Home Office to Scrap 'Racist Algorithm' for UK Visa Applicants*, *the Guardian*. Available at: http://www.theguardian.com/uk-news/2020/aug/04/home-office-to-scrap-racist-algorithm-for-uk-visa-applicants (Accessed: 11 August 2021).

Medhora, S. (2019) *Federal Court Rules Robodebt 'Unlawful' in Significant Court Case*, *triple j*. Australian Broadcasting Corporation. Available at: https://www.abc.net.au/triplej/programs/hack/federal-government-loses-major-robodebt-case/11742494 (Accessed: 12 August 2021).

Ministry of Social Development (no date) *Vulnerable Children Predictive Modelling, Ministry of Social Development, New Zealand*. Available at: https://www.msd.govt.nz/about-msd-and-our-work/publications-resources/research/predicitve-modelling/index.html (Accessed: 13 August 2021).

Murphy, S. (2018) *Confirmation Overstayers Deported Using Profiling Programme, RNZ*. Available at: https://www.rnz.co.nz/news/national/354270/confirmation-overstayers-deported-using-profiling-programme (Accessed: 12 August 2021).

Murray, A. and Giammarise, K. (2020) *Pittsburgh Suspends Policing Program that Used Algorithms to Predict Crime 'Hot Spots', Pittsburgh Post-Gazette*. Available at: https://www.post-gazette.com/news/crime-courts/2020/06/23/Pittsburgh-suspends-policing-police-program-algorithms-predict-predictive-hot-spots-crime-data/stories/202006230059 (Accessed: 8 July 2021).

NCCPR Child Welfare Blog (2017) 'Los Angeles County Quietly Drops its First Child Welfare Predictive Analytics Experiment', *NCCPR Child Welfare Blog*, 9 May. Available at: https://www.nccprblog.org/2017/05/los-angeles-county-quietly-drops-its.html (Accessed: 29 June 2021).

NeoFace® Reveal: Advanced Criminal Investigation (no date) *NEC*. Available at: https://www.necam.com/AdvancedRecognitionSystems/Products/FacialRecognition/Solutions/NeoFaceReveal/ (Accessed: 8 July 2021).

Niklas, J. (2019) 'Poland: Government to Scrap Controversial Unemployment Scoring System', *AlgorithmWatch*, 16 April. Available at: https://algorithmwatch.org/en/poland-government-to-scrap-controversial-unemployment-scoring-system (Accessed: 15 July 2021).

Nikolic, I. (2020) 'French Court Rules against Facial Recognition in High Schools', *B.I.R.D.*, 28 February. Available at: https://bird.tools/french-court-rules-against-facial-recognition-in-high-schools/ (Accessed: 14 July 2021).

Nilsson, P. (2018) *First UK Police Force to Try Predictive Policing Ends Contract, Financial Times*. Available at: https://www.ft.com/content/b34b0b08-ef19-11e8-89c8-d36339d835c0 (Accessed: 15 July 2021).

Nilsson, P. (2019) 'UK Police Test if Computer can Predict Criminal Behaviour', *Financial Times*, 6 February. Available at: https://www.ft.com/content/9559efbe-2958-11e9-a5ab-ff8ef2b976c7 (Accessed: 15 July 2021).

North Tyneside Council (2019) *North Tyneside Council Report to Cabinet*.

North Tyneside Council Report to Cabinet Item 7(l) (2015).

North Tyneside Council, Revenue and Benefit Service (no date) *North Tyneside Risk Based Verification Policy*.

#NotMyDebt (2016) *The Issue*. Not My Debt. Available at: https://www.notmydebt.com.au/the-issue (Accessed: 12 August 2021).

n-tv.de (2019) *Gewerkschaft kann Aus für Precobs-Software nachvollziehen, n-tv.de*. Available at: https://www.n-tv.de/regionales/baden-wuerttemberg/Gewerkschaft-kann-Aus-fuer-Precobs-Software-nachvollziehen-article21250215.html (Accessed: 14 July 2021).

Office of the Auditor General (2016) *Michigan IntegratedData Automated System (MiDAS)*. Performance Audit Report 641-0593–15.

Office of the Inspector General (2019) *Review of Selected Los Angeles Police Department Data-Driven Policing Strategies*. Los Angeles: Los Angeles Police Commission, p. 51.

Office of the Privacy Commissioner of Canada (2020a) *Clearview AI Ceases Offering its Facial Recognition Technology in Canada*, *Office of the Privacy Commissioner of Canada*. Available at: https://www.priv.gc.ca/en/opc-news/news-and-announcements/2020/nr-c_200706/ (Accessed: 8 July 2021).

Office of the Privacy Commissioner of Canada (2020b) *Commissioners Launch Joint Investigation Into Clearview AI Amid Growing Concerns Over Use Of Facial Recognition Technology*, *Office of the Privacy Commissioner of Canada*. Available at: https://priv.gc.ca/en/opc-news/news-and-announcements/2020/an_200221/ (Accessed: 8 July 2021).

Office of the Privacy Commissioner of Canada (2020c) *OPC Launches Investigation Into RCMP's Use of Facial Recognition Technology*, *Office of the Privacy Commissioner of Canada*. Available at: https://priv.gc.ca/en/opc-news/news-and-announcements/2020/an_200228/ (Accessed: 8 July 2021).

Oswald, M. *et al.* (2017) *Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and 'Experimental' Proportionality*. SSRN Scholarly Paper ID 3029345. Rochester: Social Science Research Network. doi: 10.2139/ssrn.3029345.

Pascu, L. (2020) *French High Court Rules Against Biometric Facial Recognition Use in High Schools*, *Biometric Update*. Available at: https://www.biometricupdate.com/202002/french-high-court-rules-against-biometric-facial-recognition-use-in-high-schools (Accessed: 14 July 2021).

Patton, J. (2018) *What Is 'ShotSpotter'? Controversial Gunshot Detector Technology Approved By Toronto Police*, *Globalnews.ca*. Available at: https://globalnews.ca/news/4344093/controversial-gunshot-detector-shotspotter-toronto-police/ (Accessed: 8 July 2021).

Pennington, P. (2020a) *Government Facial Recognition Tech Deal Offers Wide Access*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/428103/government-facial-recognition-tech-deal-offers-wide-access (Accessed: 13 August 2021).

Pennington, P. (2020b) *Police Setting Up $9m Facial Recognition System Which Can Identify People From CCTV Feed*, *RNZ*. Available at: https://www.rnz.co.nz/news/national/424845/police-setting-up-9m-facial-recognition-system-which-can-identify-people-from-cctv-feed (Accessed: 13 August 2021).

Pett, H. and Cosier, C. (2017) 'We're All Talking About Centrelink, But What Is "Robodebt" Anyway?', *ABC News*, 2 March. Available at: https://www.abc.net.au/news/2017-03-03/centrelink-debt-controversy-what-is-robodebt/8317764 (Accessed: 12 August 2021).

Pitt Cyber (2020) *Pittsburgh Task Force on Public Algorithms*, *Institute for Cyber Law, Policy, and Security*. Available at: https://www.cyber.pitt.edu/algorithms (Accessed: 8 July 2021).

Porter, J. (2020) *UK Ditches Exam Results Generated by Biased Algorithm After Student Protests*, *The Verge*. Available at: https://www.theverge.com/2020/8/17/21372045/uk-a-level-results-algorithm-biased-coronavirus-covid-19-pandemic-university-applications (Accessed: 15 July 2021).

Predictive Policing Evaluation of the Pilot Project P4 in Baden-Württemberg/Germany (2018) *Max Planck Institute for the Study of Crime, Security and Law*. Available at: http://link.springer.com/10.1007/s41125-018-0033-0 (Accessed: 14 July 2021).

Privacy Commissioner, New Zealand (2015) *Predicting Risk Without Turning Children Into 'Lab Rats', Privacy Commissioner*. Available at: https://www.privacy.org.nz/blog/predictive-risk-modelling/ (Accessed: 13 August 2021).

Privacy International (2020) *The SyRI Case: A Landmark Ruling for Benefits Claimants Around the World, Privacy International*. Available at: http://privacyinternational.org/news-analysis/3363/syri-case-landmark-ruling-benefits-claimants-around-world (Accessed: 14 July 2021).

Robertson, K., Khoo, C. and Song, Y. (2020) *To Surveil and Predict: A Human Rights Analysis of Algorithmic Policing in Canada, The Citizen Lab*. Available at: https://citizenlab.ca/2020/09/to-surveil-and-predict-a-human-rights-analysis-of-algorithmic-policing-in-canada/ (Accessed: 8 July 2021).

Scharfenberg, D. (2018) *Computers can Solve Your Problem. You May Not Like the Answer., BostonGlobe.com*. Available at: https://apps.bostonglobe.com/ideas/graphics/2018/09/equity-machine (Accessed: 7 July 2021).

Senate of Canada (2018) *The Phoenix Pay Problem: Working Toward a Solution*. Ottawa: Standing Senate Committee on National Finance. Available at: https://sencanada.ca/en/info-page/parl-42-1/nffn-phoenix-pay-problem/ (Accessed: 8 July 2021).

Settlement and Full and Final Release Agreement (2017). State of Texas County of Harris.

Sheridan, E. (2019) 'Town Hall Drops Pilot Programme Profiling Families Without Their Knowledge', *Hackney Citizen*, 30 October. Available at: https://www.hackneycitizen.co.uk/2019/10/30/town-hall-drops-pilot-programme-profiling-families-without-their-knowledge/ (Accessed: 15 July 2021).

Simonite, T. (2020) *Algorithms Were Supposed to Fix the Bail System. They Haven't, Wired*. Available at: https://www.wired.com/story/algorithms-supposed-fix-bail-system-they-havent/ (Accessed: 8 July 2021).

SRF (2017) *Solothurner Polizei will keine Software zur Einbruchsbekämpfung, Schweizer Radio und Fernsehen (SRF)*. Available at: https://www.srf.ch/news/regional/aargau-solothurn/aargauer-solothurner-polizei-solothurner-polizei-will-keine-software-zur-einbruchsbekaempfung (Accessed: 15 July 2021).

STA (2019) *Automatic Licence Plate Recognition Banned in Slovenia, Total Slovenia News*. Available at: https://www.total-slovenia-news.com/politics/4126-automatic-licence-plate-recognition-banned-in-slovenia (Accessed: 15 July 2021).

STA (2020) *Interview: Police Chief on Cybercrime, Dark Web, Drugs & Illegal Migration, Total Slovenia News*. Available at: https://www.total-slovenia-news.com/politics/5291-interview-police-chief-on-cybercrime-dark-web-drugs-illegal-migration (Accessed: 15 July 2021).

Statewatch (2020) *UK: Legal Action Threatened over Algorithm Used to Grade Teenagers' Exams, Statewatch*. Available at: https://www.statewatch.org/news/2020/august/uk-legal-action-threatened-over-algorithm-used-to-grade-teenagers-exams/ (Accessed: 15 July 2021).

Stone, L. (2020) *UK Police Scrap AI Prototype Designed to Predict Serious Violence, AIBotics*. Available at: https://www.aibotics.tech/post/uk-police-scrap-ai-prototype-designed-to-predict-serious-violence (Accessed: 12 August 2021).

Stop LAPD Spying Coalition (2018) *Before the Bullet Hits the Body – Dismantling Predictive Policing in Los Angeles*. Los Angeles: Stop LAPD Spying Coalition. Available at: https://stoplapdspying.org/before-the-bullet-hits-the-body-dismantling-predictive-policing-in-los-angeles/ (Accessed: 7 July 2021).

Tay, L. (2011) *Integration Issues Halt Vic Police's LINK Project, iTnews*. Available at: https://www.itnews.com.au/news/integration-issues-halt-vic-polices-link-project-255133 (Accessed: 12 August 2021).

tellerreport.com (2019) *Belgian Police Stop Facial Recognition At Zaventem Airport, Tellerreport.com*. Available at: https://www.tellerreport.com/tech/2019-09-21---belgian-police-stop-facial-recognition-at-zaventem-airport-.BkEeQN8QDH.html (Accessed: 8 July 2021).

Tunney, C. (2020) *RCMP Says It Will Limit Its Use of Facial Recognition Tech — But Won't Stop Using It Entirely, CBC*. Available at: https://www.cbc.ca/news/politics/rcmp-clearview-ai-1.5490988 (Accessed: 8 July 2021).

United Nations Special Rapporteur (2020) 'Brief by the United Nations Special Rapporteur on extreme poverty and human rights as Amicus Curiae in the case of NJCM C.s./De Staat der Nederlanden (SyRI) before the District Court of The Hague (case number: C/09/550982/HA ZA 18/388)'.

University of Cambridge (2018) *Helping Police Make Custody Decisions Using Artificial Intelligence, University of Cambridge*. Available at: https://www.cam.ac.uk/research/features/helping-police-make-custody-decisions-using-artificial-intelligence (Accessed: 15 July 2021).

University of Copenhagen (2019) *New Technology to Prevent Future IT Flops in the Public Sector, Faculty of Science, University of Copenhagen*. University of Copenhagen. Available at: https://science.ku.dk/english/press/news/2019/new-technology-to-prevent-future-it-flops-in-the-public-sector/ (Accessed: 14 July 2021).

Victorian Government Health Information (2012) *HealthSMART - Victorian Health IT*. Available at: https://web.archive.org/web/20120618013645/http:/www.health.vic.gov.au/healthsmart/ (Accessed: 12 August 2021).

Wakefield, J. (2020) 'A-Levels: Ofqual's "Cheating" Algorithm Under Review', *BBC News*, 20 August. Available at: https://www.bbc.com/news/technology-53836453 (Accessed: 15 July 2021).

Webb, S. (2017) *Houston Teachers to Pursue Lawsuit over Secret Evaluation System, Houston Chronicle*. Available at: https://www.houstonchronicle.com/news/houston-texas/houston/article/Houston-teachers-to-pursue-lawsuit-over-secret-11139692.php (Accessed: 29 June 2021).

Why Do Projects Fail? (2012) 'HealthSMART – Victoria', *Why Do Projects Fail?*, 11 November. Available at: http://calleam.com/WTPF/?p=3847 (Accessed: 12 August 2021).

Wikipedia (2021) 'PRECOBS', *Wikipedia*. Available at: https://de.wikipedia.org/w/index.php?title=PRECOBS&oldid=210617276 (Accessed: 15 July 2021).

Wills, T. (2019) 'Sweden: Rogue Algorithm Stops Welfare Payments for up to 70,000 Unemployed', *AlgorithmWatch*, 25 February. Available at: https://algorithmwatch.org/en/rogue-algorithm-in-sweden-stops-welfare-payments (Accessed: 15 July 2021).

Wisterberg, E. (2018a) *12 av 16 säger upp sig i protest mot kommunens bidragsrobot*, *Breakit*. Available at: https://www.breakit.se/artikel/11082 (Accessed: 15 July 2021).

Wisterberg, E. (2018b) *Titta! Kommunens robot avgör vem som får socialbidrag…*, *Breakit*. Available at: https://www.breakit.se/artikel/10770 (Accessed: 15 July 2021).

Yogaretnam, S. (2020) Chief Says Facial Recognition Software Test Drew on Police Mug Shot Database, Ottawa Citizen. Available at: https://ottawacitizen.com/news/local-news/chief-says-facial-recognition-software-test-drew-on-police-mug-shot-database (Accessed: 8 July 2021).

Carnegie UK
Andrew Carnegie House
Pittencrieff Street
Dunfermline
Fife, Scotland
KY12 8AW

T +44 (0)1383 721445

Carnegie United Kingdom Trust
Registered Charity No: SC 012799 operating in the UK
Registered Charity No: 20142957 operating in Ireland
Incorporated by Royal Charter 1917

collective wellbeing
Carnegie UK