CURATED PROCEEDINGS

ADM +S ARC Centre of Excellence for **Automated Decision-Making and Society**

ADM+S **WORKING PAPER SERIES**

ADM in child and family services
Mapping what is happening and what we know

May 2022

Editors
Dr Lyndal Sleep
Prof Paul Henman

WORKING PAPER 002

admscentre.org.au/publications

ACKNOWLEDGEMENT OF COUNTRY

In the spirit of reconciliation, we acknowledge the Traditional Custodians of country throughout Australia and their connections to land, sea and community. We pay our respect to their elders past and present and extend that respect to all Aboriginal and Torres Strait Islander peoples today.

ABSTRACT

This discussion paper reports on the proceedings of a workshop entitled ADM in child and family services: Mapping what is happening and what we know on 24 November 2020. This workshop brought together international experts and on-the-ground sector stakeholders to provide an overview of what automated decision-making is being used in child and family services and what it means to service providers, professionals and service users. Philip Gillingham presented on ADM in child and family services in Australia and beyond, Joanna Redden on her research mapping ADM in Child and Family services in the UK. Rhema Vaithianathan's presented on her work designing predictive analytics systems for the child and family services sector in the United States. Carol Ronken from advocacy organisation Brave Hearts talked about both the opportunities for ADM to enhancing child safety, as well as its challenges. The presentations were followed by a round table discussion that is summarised in the report. The workshop and this discussion paper aim to provide baseline knowledge for further research in this emerging and rapidly changing area.

KEYWORDS: AUTOMATED DECISION-MAKING (ADM), CHILD AND FAMILY, CHILD PROTECTION, CHILD SAFETY, SOCIAL SERVICES, PREDICTIVE AI.

Suggested citation: Sleep, L. and Henman, P. (2022) *ADM in child and family services: Mapping what is happening and what we know*, ADM+S Working Paper Series 2022 (002), ARC Centre of Excellence for Automated Decision-Making and Society, DOI: 10.25916/n86m-km33

TABLE OF CONTENTS

1. INTRODUCTION

Prof Paul Henman

The purpose of this discussion paper is to report the proceedings of a workshop on automated decision-making in child and family services which was conducted as part of the ARC Centre of Excellence for Automated Decision-Making and Society's (ADM+S) Social Services Workshop series. The workshop was hosted by the University of Queensland on 24 November 2020. This workshop aimed to bring international experts together to provide an overview of what automated decision-making is happening in child and family services and what it means in the sector. It is hoped that this discussion paper provides baseline knowledge for further research in this emerging and rapidly changing area.

This discussion paper is organised into four main sections, which reflect the proceedings of the workshop. The first section provides an introduction to the ADM in Social Services research conducted by Paul Henman, who hosted the workshop at the University of Queensland, and offers definitions for the terms ADM and a brief outline of Child and Family services. This is followed by summaries of each of the keynote presentations of cutting-edge research on ADM in child and family services. A summary of Philip Gillingham's presentation of ADM in child and family services in Australia and beyond, drawing on Philip's work on Structured Decision-Making in the sector, is followed by Joanna Redden's research presentation on mapping ADM in Child and Family services in the UK. Finally, Rhema Vaithianathan's presentation regarding her work designing predictive analytics systems for the child and family services sector in the United States is summarised. The next substantive section of the report is a summary of Carol Ronken's presentation detailing her observations from the field of child and family service practice in Australia. The final main section of the report summarises the round table discussion that followed the presentations in the workshop. In this section, some key topics/concerns that were discussed are presented in the form of transcripts and organised under key headings with the intention of providing an indication of key research needs for a future research agenda.

Watch the event recording on YouTube.

About the ADM+S Centre

The ARC Centre of Excellence for Automated Decision-Making and Society (ADM+S) brings together universities, industry, government and the community to support the development of responsible, ethical and inclusive automated decision-making. The Centre has identified a number of substantive areas through which to focus our research over the seven years. One of those primary areas of focus is that of social services. In its first twelve months, the Centre will conduct several workshops with this being the first of the series. The workshops have been designed to build a knowledge base that will inform the Centre's research programs, engagement and impact. The social services focus areas that will be included beyond this particular event are:

- social security and income support (held 5 May 2021)
- disability services (held 20 September 2021)
- migration services
- criminal justice

What does social child and family services mean and what does automated decision-making mean?

Child and family services covers a wide range of services. Typically, child protection focuses on identification and management of child abuse and neglect and includes out of home care, parenting education and training and family support. The sector has human service professionals contributing both as state and non-state actors. Child and family services are a State (regional) responsibility in Australia. In addition to child protection, Australian governments have increasingly created advocacy services such as Children's Commissioners and children's rights agencies.

Automated decision-making (ADM) is the use of digital technologies and algorithms to automate all or parts of human decision-making. In areas of child and family services risk assessment has been a key part of professional practice in which ADM has been increasingly deployed. Risk assessment is very much focused on decision support skills tools, but there are other areas where we believe automated decision-making may become helpful in the area of child and family services. These other areas where the implementation of ADM may be beneficial are canvassed and critically examined within this workshop.

2. KEYNOTE PRESENTATIONS

ADM in Child and Family Services: Mapping what is happening and what we know

Dr Philip Gillingham, The University of Queensland

My interest in ADM

My PhD research, completed in the early 2000s, focused on how Child Safety Officers (CSO) in Queensland employed Structured Decision-Making (SDM) tools and found that they were not being used as intended. The SDM tools were embedded in an electronic information system (IS) and so, more broadly, I also conducted research about how CSOs used the information system.

From 2013-2015, I held an Australian Research Council Discovery Early Career Research Award to research how IS might be better designed to support rather than hinder social work practice. During this time, and in response to the rapid development in the field, my attention turned to how the data that was accumulating in IS was being used, particularly for predictive modelling. Since 2017, I have held an Australian Research Council Future Fellowship, to continue this work.

My research has identified some jurisdictions that are implementing, or have implemented ADM technologies, as well as some key issues that have emerged:

Where ADM is being (or touted to be) used in child and family services

Aotearoa/New Zealand

Around 2014, Aotearoa/New Zealand attempted to develop a Predictive Risk Model. However, it was discontinued and did not reach trial stage (Vaithianathan et al 2013; Gillingham 2016). A key point to derive from this example was that the category of 'substantiated cases' contained many children who had not been maltreated, hence predicting 'substantiation' is not the same as predicting 'maltreatment' (Gillingham 2016). Further, it should be noted that the main predictors identified within the PRM were the relationship status of the caregiver (single parent), the caregiver's care and protection history as a child, and the care and protection history of other children in the family (Gillingham 2017).

England

The What Works Centre for Children's Social Care worked with four local authority child protection services to test eight predictive models built using both structured and unstructured data (textual analysis). None of their models were considered to be sufficiently accurate in their predictions to be of any use in practice. (Clayton et al 2020; Denzik et al 2018)

Other approaches to ADM in child protection have sought to provide integrated databases to support child protection decision-making. Some examples are identified below.

- At Bristol Integrated Analytical Hub:

The Troubled Families programme was launched in 2011 to help families who struggle with factors such as unemployment, crime and poor school attendance. Think Family identified families facing issues such as, parents and children involved in criminal or anti-social behaviour; children not attending school regularly; children who need help; adults out of work or at risk of financial exclusion, and young people at risk of worklessness; families affected by domestic violence and abuse; parents and children with a range of health problems. (Denzik et al 2018, p.27)

- Another location was the Camden Resident Index:

The Camden Resident Index is a data management system utilising software supplied by IBM that allows for a ‘single view of a citizen’ by aggregating data from sixteen different council business systems across Camden Council, covering 123 fields of primarily demographic information..... The Camden Resident Index is used by the Multi-Agency Safeguarding Hub to locate information about a household’s engagement with services across the Council. A key use of the index is to enable fraud detection, such as validation for residency and for accessing council services such as school places, the number of residents in a household eligible for a council tax discount, or in cases of illegal subletting of council housing. (Denzik et al 2018, p. 48)

- At Manchester’s Research & Intelligence Database:

Manchester City Council is using a system it calls the Manchester Research & Intelligence Database to identify families in need of support and to enable caseworkers to access information more quickly than previously, and to enable the collection and analysis of data to assess services and impact. The Council developed the system by buying an IBM product called iBase. The approach to using the system has been developed internally and its aim is to empower ‘lead professionals to make the best use of data they are legally able to see, and is not about replacing decision-making or interpretation with system algorithms or decisions’. The stated aim sought to enable a more ‘holistic’ understanding of people, needs and services. Future goals include developing decision-making tools, building performance reporting tools, building threshold and alerting tools and rolling the system out to more users, as well as connecting more data. (Denzik et al 2018, p.26)

- Finally, in Hackney’s Early Help Profiling System:

In child welfare, Hackney County Council worked with Ernst & Young (EY) and Xantura on the use of a system to identify children at risk of maltreatment and families who need additional support. The system was called the “Early Help Profiling System” (p. 55) however it was subsequently discontinued in 2019 after a councillor alerted the media and there was public outcry regarding potential invasions of privacy.

USA

In Chicago, a Rapid Safety Feedback program was withdrawn after it started to make increasingly mistaken and bizarre predictions. The program had been developed by a private company with no public scrutiny.

In Pennsylvania, a Family Screening Tool (FST) was used in Allegheny County. However, practitioners are already second-guessing FST outcomes, and parents have shown concern about the inclusion of 'previous involvement' as a predictor of future abuse/neglect. Eubanks (2018) and Gillingham (2020) have also discussed the Allegheny Family Screening Tool. Two external evaluations have been conducted into the Allegheny Family Screening Tool (Allegheny County Department of Human Services, 2019; Hornby Zeller Associates, 2018). See also Vaithianathan et al (2017) on the development of the Tool.

In California, research into these types of programs has been carried out by Drake et al. (2018). Other useful resources include the Children's Data Network and California Department of Social Services (2020), and *Predictive Risk Modelling: Practical Considerations* (2020).

For further information on Florida, Kelly (2015) serves as a useful resource.

The Netherlands

In the Netherlands, Amrit et al. (2017) developed a decision support system to support paediatricians and other healthcare workers in Amsterdam to identify children under their care who are at risk of maltreatment. They used healthcare records, which they claimed were complete, and both structured data (entered into a specific field, like age, height and weight) and unstructured data (the notes made by healthcare workers in free-text fields), an approach they claimed was unique at the time. They assert the system is 90% accurate. As yet it has not been implemented.

How ADM is being used in child and family services (e.g. screening; decision support; resource allocation; matching; risk assessment; risk reduction)

Some of the proponents of ADM, such as Pedersen (2019), take an extreme position and argue that ADM could be used for all of the above functions. Pedersen (2019) argues that with recent advances in information science, 'relationshipism' (his limited understanding of social work) should be replaced by a model of service delivery he terms 'dataism'. Indeed, Pedersen provides a good example of the myopic debates about the use of ADM. Even if accurate and useful ADM can be developed in the future, they will only constitute a very small part of the process of delivering and evaluating services. Social workers still need to assess needs, identify threats to safety and develop and deliver individually suitable interventions (Devlieghere et al, 2021).

Legal, ethical, organisational, data challenges of applying ADM in child and family services

Legal

No research has been conducted as yet into the legality of predictive modelling. Although the legality of trawling through data to find and intervene in the lives of families, as happened in Hackney, may have led to a legal challenge if Hackney had continued this practice. Prior to Brexit, handling children's data in the UK was subject to the General Data Protection Regulation which was applied across the EU in 2018. The following document outlines their rules for dealing with data about children, but the GDPR also stipulates that "the GDPR and Data Protection Act 2018 do not prevent, or limit, the sharing of information for the purposes of keeping children and young people safe".

HM Govt. (2018) *Information sharing, Advice for practitioners providing safeguarding services to children, young people, parents and carers* (July 2018) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/721581/Information_sharing_advice_practitioners_safeguarding_services.pdf

Ethical

The main concerns about predictive modelling that have arisen are related to privacy, information sharing between agencies, and using information collected for one purpose for a different purpose (Dare 2013; Drake and Jonson-Reid 2018; Leslie et al 2020). Useful academic commentary has been done by Keddell (2015) and de Haan and Connolly (2014) in this regard. Algorithmic accountability, in terms of how agencies using ADM can explain decisions and be 'transparent', has also been raised as a challenge by Gillingham (2019), Binns (2017) and USACM (2017).

Organisational Data Challenges

The data in the form of case files has been criticised in terms of its completeness and whether or not it accurately represents the realities of social work practice. Machine learning using data created by humans and collated within administrative structures may just end up replicating the biases of previous practice. There is also the problem of using 'proxies' for child maltreatment or the risk of child maltreatment, such as re-notification and substantiation, as they may lead to tools that predict administrative processes rather than actual maltreatment. Salganik (2018) suggests that the accuracy of predictive modelling might be improved by developing 'custom' datasets which collate information about the factors and processes which research has shown are associated with the phenomenon to be predicted. See also Gillingham (2020a; 2020b; 2019; 2015a; 2015b).

Professionals and administrators' engagement with ADM

In my PhD research regarding how the Structured Decision-Making tools were used in Child Safety, I found that although it appears that the tools were used in decision-making (data had to be entered to proceed through the workflow of the information system), the metadata (backed up by interviews and observations) showed that the forms were filled in some time after decisions were made.

Workers had to go through the workflows to 'close' a case and that was when the SDM tools were engaged with. This is due to a myriad of reasons, such as static variables (factors that will never change, for example, having been an illicit drug user/suffered mental illness or involved with the police). A key theme was that the SDM tools were good at 'stating the obvious' but were not helpful in complex cases, where support was required. See Gillingham (2011; 2009a; 2009b) and Gillingham and Humphreys (2010) on these points.

My PhD research provides possible hints as to how ADM in child welfare might be regarded by workers. However, the research was completed 13 years ago and further research is required to account for changed contexts, cultures and ADM. Nevertheless, the Chouldechova et al. (2018) article begins to shed some light in this area reporting that 25% of recommendations by the Allegheny *Family Screening Tool* are over-ridden.

Users' understanding and experience of processes that involve the use of ADM

With the exception of Eubanks' (2018) chapter, there is little research about how services users in child welfare understand and experience ADM.[1] Problems may arise when citizens demand to know how ADM affects decision-making.

Data and research knowledge used to develop ADM

As outlined above, it is mainly administrative data that are used to develop ADM. Research regarding the etiology of child maltreatment, for example, is rarely considered. One of the reasons for this is that, from an information science perspective, it does not matter what researchers might find because the real 'answer' is in the data. The correlates that a machine learning exercise identifies within a large set of data is given exalted authority. This tendency is linked to debates about the end of 'theory' in the context of big data. However, the administrative data from welfare agencies is often incomplete and provides a solely administrative account of assessment and service delivery. At most, administrative data only provides a very partial account of what happens, for example, in cases of child maltreatment.

One of the key challenges in developing ADM is identifying what factors are predictive (causal rather than just correlated) and how they can be best captured, categorised, and analysed.

[1] Exceptions include, for example, impact evaluation and process evaluation for AFST and RCT results for Douglas County Decision Aid https://csda.aut.ac.nz/__data/assets/pdf_file/0012/504102/Douglas-RCT-Final-Report-210211.pdf

Emerging thoughts... and potential controversy

In the world of information science, software is often referred to as a 'solution'. In order for it to be a 'solution', it follows that there was a problem to be solved in the first place. With ADM in child protection services, the 'problem' is conceived as decision-making, including risk assessment, which research has demonstrated can be subjective, biased and inconsistent. ADM is promoted as the solution which will address these problems and improve decision-making. The SDM tools that were introduced in Queensland for this very purpose made no discernible difference to child protection services. Within a few years, the service was considered to be 'in crisis' again, to the extent that it was deemed necessary to conduct yet another lengthy and expensive review followed by a re-organisation.

While I agree that decision-making in social work can always be improved, a recent project about children killed by their parents reminded me that, most of the time, social workers are capable of making good decisions and can continue to do so as circumstances change and new information comes to light. The inconsistency of decision-making in social work has frequently been invoked as a reason to develop decision-making tools. Inconsistency is usually identified by looking at patterns in the data about service delivery. However, taking a closer look at what social workers actually do, and in particular the context in which they do it, reveals that what appears as inconsistency from afar is, up close, pragmatic decision-making.

ADM in social work is in its infancy and thus far there has been little success. Considerable investment is required to develop ADM to the point where useful tools can possibly be produced. A key question to consider therefore, is whether such an investment would lead to commensurate improvements in service delivery. To paraphrase Hollnagel and Woods (2005), as a new form of technology, ADM would have to meet either or both of the following criteria: firstly, ADM would need to improve work performance significantly (better decision-making) and secondly, ADM would need to make possible some task or function that cannot currently be done (mapping of future service needs).

Provocations

When researching emerging forms of ADM in social services, the following questions should be perpetually emphasised.

- Why do we need ADM in child protection services?
- Will ADM ever be as accurate as practitioners?
- Does ADM oversimplify the complexities involved in dealing with child maltreatment?
- Will ADM simply 'state the obvious' and therefore can it be helpful in extreme cases?
- Irrespective of how ADM might be used, decision-making is only a very small part of dealing with child maltreatment. Is ADM worth the investment?

Datafied Child Welfare Services: Politics and Policy

Assistant Prof Joanna Redden, University of Western Ontario (Canada)

It was a pleasure taking part in this event, hosted by the ARC Centre, to map what is happening and what we know about the use of automated decision-making systems in child and family services. At the Data Justice Lab (datajusticelab.org), we have been mapping and analysing government uses of data systems with a focus on uses of predictive analytics. In this short summary I draw on research that has been published in our 'Data Scores as Governance' report (Dencik et al. 2018) as well as research that discusses the political implications of uses of predictive analytics in child welfare in the UK (Redden et al. 2020). Our research in this area has involved the use of multiple methods. In short, these methods include: scraping relevant documents from government and media sites; multi-stakeholder workshops; freedom of information requests (423); six case study investigations; and interviews with system developers, practitioners and civil society groups. We built the data scores tool to make our information publicly accessible and to enable further research (data-scores.org).

In the UK, predictive data systems have been trialled or remain in use in Thurrock, Newham, Tower Hamlets, Somerset, Hackney and Bristol. The What Works Centre for Social Care partnered with the Office of the Children's Commissioner to research risk scoring systems in four councils that have not been named. We are also aware of another council that piloted use of predictive analytics for child welfare and decided not to proceed with using the system.

Our research shows that in relation to child welfare, different UK local authorities have taken different approaches to the use of data technologies. Three of our six case study investigations involved local authorities' use of data systems in the area of child welfare, only two of these systems involved predictive risk assessment.

The Bristol Integrated Analytical Hub is a system developed in-house that makes use of a database that consolidates 35 social issue datasets from about 54,000 families. Initially the Hub was created as a 'data warehouse' in response to the Troubled Families programme. After the data warehouse was developed the Bristol team began looking into ways to predict future needs and created a model for predicting child sexual exploitation, developing a risk score for all young people in Bristol.

Hackney trialled the Early Help Profiling System through a partnership with Ernst and Young and a company called Xantura. This System was used to identify children at risk of abuse or neglect. The system used 'a predictive risk model which brings together data from multiple agencies'. The system used a model that analyses data so that monthly risk profiles are sent to social workers for those families identified as most in need of intervention.

Manchester purchased an IBM system called iBase and then modified and developed it according to their needs. Their system is called the Manchester Research & Intelligence Database and it is used to identify 'troubled families' and help caseworkers working with these families. The aim of the system according to developers is to make it possible for workers 'to make the best use of data they are legally able to see'. As of 2016, the data warehouse that Manchester created combines 16 datasets, and caseworkers are able to access data going back five years. As with Bristol, the stated aim is to enable a more 'holistic' understanding of people, needs and services. Unlike Bristol and Hackney, the Manchester system did not do family level prediction at the time of our research. Although it was noted that future goals include developing decision-making tools and alerting tools, it is not clear if these will perform risk assessments.

In our research we discuss the political and policy implications of these systems by using Kitchin and Lauriault's data assemblage as a framework of analysis (Kitchin 2014, Kitchin and Lauriault 2014). We argue that our analysis demonstrates why these systems must not be considered neutral decision aid tools and should instead be viewed as situated technologies tied to complex assemblages of people, practices, policies, politics, and technical artefacts. We find that local authorities are turning to these systems largely because of the austerity program introduced by the Conservative government, post financial crash, which saw some local authorities see their budgets cut by as much as a third. As a result, local authorities are trying to help more people with less. The systems are viewed as a way to provide more targeted services. The austerity program itself reinforces a long-standing neoliberal logic which promotes reduced state services. The use of these systems to target resources to 'most need' cases in practice reinforces the idea that there needs to be a narrowing of support provided. The use of 'risk assessment and scoring' data tools also further embeds a logic of citizens as risks and an emphasis on individual responsibility rather than a focus on wider social economic causes of family crisis.

The British Troubled Families program is a driving force. This policy initiative compels local authorities to identify families that meet a number of criteria in order to be labelled 'troubled'. Once labelled there is an ability to access more resources. The policy and program compel local authorities to link up datasets about people. The program itself has been widely criticised for individualising social problems and stigmatising families. Concerns are being raised about a lack of attention paid to the impact of extensive data collection and sharing on people subject to these systems. Rights organisations are raising concerns about implications of this kind of labelling in a datafied landscape when it becomes easy for labels to stick with people and become amplified. Legally, local authorities' duty of care responsibility is the main justification being used for predictive risk assessments, but this is being challenged by some who argue for greater debate about the need for rights protection particularly in light of limited to no impact assessments. We also find local authorities taking different approaches to consent for data use, with some seeking consent to use people's data and others not doing so.

There is a growing data marketplace that requires attention. Each local authority analysed engaged differently with the larger data marketplace. Our research also identified differing attitudes about the appropriate role for private tech companies in the area of social care that has consequences for differing levels of transparency.

In the area of predictive analytics, we found a lack of information available about how the use of these systems affect resource allocations and decision makers. We raise concerns about the lack of investigation into the impact of these systems on families and children. Engaging with these questions will become more pressing as the data marketplace for child welfare and social care grows. We found differing levels of transparency, accountability and oversight which also requires greater attention.

The use of data lakes and predictive data systems need to be considered as part of ongoing efforts to modernise and rationalise the way social workers collect information and make decisions through the introduction of various computational technologies over the last number of decades. Previous research has highlighted how these information systems can limit ways of knowing (Gillingham 2015a, Keddell 2015, Munroe 2010, White et al. 2010) which in turn limits the potential of this data to be used for prediction. We find the turn to predictive technologies in social care as a continuously contested and negotiated space. Concerns are being raised about how uses of this technology may transform social work, increase surveillance, entrench hierarchies of knowledge and an emphasis on quantification over relationships. Since our report was published, research by the What Works for Children's Social Care Centre has raised concerns about the accuracy and effectiveness of predictive technologies in social care (Clayton et al. 2020).

In conclusion, our research points to the political and economic factors influencing the development of data systems in social care and why these systems must not be considered neutral decision aid tools. We detail how systems of thought, ownership structures, policy agendas, organisational practices and legal frameworks influence the way data systems are developed. There is a general trend toward greater data sharing in social care but also differences in perspectives about rights and the value of predictive technologies which demonstrates the need for greater and more widespread and inclusive debate about the use of predictive systems. There is far too little attention given to the need to investigate the impact (both intended and unintended) of these systems on families. This research points to the need for policy makers to engage with fundamental questions of power dynamics, rights, and participation.

Developing ADM technologies in Child and Family Services

Prof Rhema Vaithianathan, Auckland University of Technology

Professor Rhema Vaithianathan's contribution to this discussion of the use of automated decision-making (ADM) and machine learning models in child and family services, is informed by her work in developing and implementing decision support tools with a number of government agencies operating in this complex area. This work has been developed with researchers at the Centre for Data Analytics (CSDA), at Auckland University of Technology (NZ) and The University of Queensland.

CSDA research context: Automated Decision-Making vs Decision Support Tools

Researchers at CSDA have developed several predictive analytics tools that are being used by governmental agencies — including departments in Allegheny County and Douglas County in the US — to help inform decisions about child welfare, family support and homelessness programmes.

Critically, it should be noted that these tools have not been designed to automate or replace decision-making by frontline workers in these agencies. The intended function of these tools is to make use of available administrative data and provide additional timely and meaningful information to frontline workers to support better, more informed decisions.

Using this approach, CSDA researchers are developing tools that support timely, quality decision-making in complex areas, such as social services, while retaining "the human in the loop".

CSDA tools use predictive risk modelling (PRM) which gathers and collates data from a range of sources (such as child welfare and public welfare eligibility systems) and then processes this data to generate a screening score.

Making sense of available data: How Predictive Risk Models are being used in child and family services

Two of the US counties that CSDA researchers are working with are using PRM tools and the resulting screening — or risk — score to help triage calls relating to child maltreatment and neglect. The significance of this sort of usage to support decision-making in child and family services is that these tools can process a lot of information quickly.

Often these frontline workers are faced with making a decision about a child who may have already had multiple referrals or removals and who may have a complex family and case history, and therefore the frontline worker may struggle with information overload. This information overload is paired with pressure to make a decision in a relatively short space of time: in triaging these calls the average time spent determining whether or not to screen in and investigate a child and family is 10 minutes.

Under time pressure and facing a massive amount of data to process, the risk is that the frontline worker cannot practically make use of the available information. This is a concern shared by a number of the agencies that CSDA works with.

As Brian Taylor, Professor of Social Work at Ulster University, has noted: “[Child welfare] organisations emphasise gathering all the relevant information … what use is to be made of the information is a matter for professional discretion.” (Taylor, 2017)

PRM tools can rapidly process large amounts of data and use this information to generate a useful indicator of the risk of maltreatment and neglect. They can also quickly provide an indication of the case complexity, alerting frontline staff to the fact a complex case history exists and that more investigation needs to be carried out. Significantly, PRM tools summarise a large amount of data into a single index, enabling people to access and apply the information to support their decision-making process.

Without these tools, the default position for frontline staff facing information overload and time pressures may be to apply heuristics or mental shortcuts to help make faster decisions. These shortcuts can increase the risk for unconscious bias. Stanford University’s Professor Jennifer Eberhardt (Eberhardt, 2020) has identified greater risk for unconscious bias in situations where people are overwhelmed or pressured and are faced with making decisions in situations where there is ambiguity and complexity.

The other driver for the development and use of PRM tools is that child welfare systems are themselves under pressure. Data from the US shows that 1 in 3 American children will be investigated for abuse and neglect before they turn 18. For African American and Black children in the US, that statistic is 1 in 2. Not all children who are reported are in need of intervention but determining which children do require intervention is complex.

The US child welfare system was designed to identify and respond to serious abuse and neglect, but it is touching the lives of a third to half of all families. One of the reasons for this expansion of the system — and arguably a lack of appropriate focus of resources — is because up until now we have provided very few tools for people to make better screening decisions.

What we do with PRM is training a model using features like prior child welfare and public benefits history and combining them using a statistical method to create a score, which is validated and correlated to that particular population. So, unlike a lot of the actuarial and structured decision-making tools that are available, we build these PRMs for the population they will serve, using local data.

Accuracy and bias: What are the legal, ethical, organisational, data challenges of applying ADM in child and family services?

CSDA's PRM tools in use by frontline workers to support decision-making in the triaging of calls relating to child maltreatment and neglect, deliver a score of between 1 and 20 to indicate a child's risk of removal from the home within the next two years. A score of 1 indicates a low risk and 20 indicates the highest risk of removal. Some of the key concerns that have been raised about these scores are that too much weight could be placed on an individual score and that the score will not always provide an accurate indication of risk.

Before discussing accuracy and bias in predictive risk models in use in child and family services, it is necessary to discuss the bias issues that already exist in the delivery of these services.

In almost every county we work with, we see a racial disparity in how children with similar risk (according to retrospective screening scores) are being treated. In the US, we have been able to go back and look at the data and the screening decisions that were made about children and see that a child with a similar risk profile who is Black is being treated or investigated more commonly than a child who is White with the same risk profile.

Machine learning can provide a more objective, data-driven decision, although it must be acknowledged the data a PRM tool uses may contain some inherent bias (i.e., surveillance bias may lead to a greater number of interventions).

With our PRM tools we acknowledge the scores they generate are not always 100% accurate, but they provide a useful prediction of a child's risk and they support better screening decisions that more closely correlate with outcomes. So, with our PRM tool results, it should be noted that only 50% of the children who scored 20 go on to be removed in the next two years, and there are some children who score a 1 and go on to be removed.

However, when we go back and look at historical screening decisions and retrospectively generate risk scores, we see that agencies were commonly screening in lots of children who had very minimal chance of having any chronic involvement with child welfare services and screening out a lot of children with very high risk.

In terms of implementing ethical checks and balances, CSDA is very committed to putting these in place and has a keen focus on transparency. Drawing on other work in this area we have created the "CSDA guardrails for PRM projects" which outline six key principles we have adopted around the development and implementation of these models.

human centered data science
guardrails for PRM projects

- Agency leadership
- Multi-disciplinary team
- Transparency and fairness
- Ethical review
- Community voice
- Independent Evaluation

Agency leadership is critical. We believe the agencies should be in the driving seat with these tools, and even though these models can be considered complex and 'black box', the teams purchasing them should require the vendors (the developers of the tools) to explain how they work.

Transparency around how these tools work is very important and CSDA publishes all its methodology documents online.

We also focus on building multidisciplinary teams. We have people working in ethics and fairness and disparities evaluation, and we build in an evaluation component that is independent of our group to provide additional checks and balances. The agencies we work with often commission independent ethics reports on our work (e.g. *Allegheny Family Screening Tool: Ethical Analysis PDF*)

One of the greatest concerns we have around ethics and fairness in child welfare is that of surveillance bias. In child welfare data systems we do not have ground truth of child maltreatment, we have what is reported into our child welfare systems. A child who is at no greater risk of neglect or maltreatment than another child may be reported to child welfare multiple times for multiple reasons. To counter this, we try to look for more objective "ground truth" measures of adversity against which to validate our tools. Actual incidents of fatality or near death can provide this ground truth universal measure of adversity because these incidents are usually reviewed by people outside the child welfare system, and they are subject to a much more rigorous analysis.

Community voice and consultation is also an important principle in our work. The agencies we work with lead this community engagement, but it is a critical element of our work, which draws on participatory design and looks to engage with, and address concerns around, the use of data in complex settings such as child and family services.

Some CSDA projects applying Predictive Risk Modelling tools in child and family services

Allegheny Family Screening Tool

Researchers led by Vaithianathan modelled, designed and supported implementation of this world-first child welfare predictive analytics tool. The Allegheny Family Screening Tool (AFST) uses rich administrative data to generate a screening score for incoming calls alleging child maltreatment and neglect. The score is an additional piece of information that helps call screeners as they decide whether to open an investigation. Allegheny County introduced this decision support tool with the aim of improving accuracy and consistency of call screening decisions. An independent impact evaluation of the tool was completed by researchers at Stanford University in March 2019. The evaluators' findings included that use of the tool improved the accurate identification of children in need of services and was associated with a modest reduction in racial disparities in case openings.

Read key Allegheny Family Screening Tool documents (including Methodology, FAQs, Process Evaluation and Impact Evaluation) as a PDF.

Allegheny County's Hello Baby Program

Allegheny County's Hello Baby Program is intended to provide every family of a new-born with universal access to information and provide resources and differentiated and intensive support for families with complex challenges and needs. The CSDA research team was engaged by Allegheny County Department of Human Services to develop a screening tool that would allow for identification of families with highest needs. The agency will contract externally for programmatic intervention and independent evaluation.

Read key Hello Baby documents (including Methodology, FAQs and Ethics Reports) on the Allegheny County Analytics website.

Douglas County Decision Aid

In early 2017 researchers developed a prototype child welfare (maltreatment) predictive risk model for Douglas County, Colorado. The Douglas County Decision Aid is a decision support tool designed to help triage incoming calls alleging child maltreatment and neglect. The tool uses data from child welfare and public welfare eligibility systems. The DCDA was implemented by Douglas County leadership, initially as a year-long randomised controlled trial, in February 2019, with results of the impact evaluation published in 2021.

Read key Douglas County Decision Aid documents (including Methodology and Impact Evaluation) on the Centre for Social Data Analytics website.

Observations from the field

Carol Ronken, Bravehearts

Child protection workers typically have large caseloads that require engaging with multiple sources to gather information and are under pressure to make assessments quickly. One of the most difficult tasks in child protection is assessing risks, particularly in relation to appropriate responses to alleged concerns.

As an organisation that works with children and young people who have been sexually harmed, we are all too aware of the difficulties in identifying children at risk, as well as those who are being or have been harmed. Statistics vary depending on the source of the figures, but most agree that approximately one in five children will experience some form of sexual harm before their 18th birthday (Price-Robertson, Bromfield and Vassallo, 2010; Mills, Kisely, Alati, Strathearn, & Najman, 2016; Royal Commission into Institutional Responses to Child Sexual Abuse, 2017). Being able to identify children vulnerable to, or experiencing sexual harm, as well as other forms of abuse and neglect, is essential in ensuring adequate, effective, and timely interventions.

Early identification and intervention for children and young people vulnerable to harm has the potential to both reduce the risk of long-term adverse outcomes for children and to support child protection workers and police to better respond to notifications. Improving systems to accurately identify vulnerable children is critical for policy makers, child protection and police, as well as child protection advocates and researchers, allowing for targeted and effective allocation of child protection resources.

Benefits of an automated decision-making approach in the child protection arena is that it may be able to identify patterns and correlations more readily and more effectively across information collated from various sources, and more objectively assess data profiles for risk, reducing the potential for erroneous or biased human assessment.

Considerations

Data Sources

There are many barriers to victims of sexual offences disclosing and seeking support. Incidence studies, which measure the number of new cases occurring during a one-year period, reflect only cases that are officially reported to authorities and fail to recognise the many cases (estimated at 95 - 97%) that go unreported (Martin & Silverstone, 2013). Additionally, in respect to child sexual offences a large majority of these do not fit within the remit of child protection departments as the child's home environment is often safe. While the majority of child sexual offences occur within familiar and familial relationships (Quadara et al., 2015), the perpetrator is often outside of the child's immediate home (Tarczon & Quadara, 2012).

In late 2019, the Australian Government announced an investment in rolling out iTree's child safety intelligence system, **REACH**. The system has been refined to allow state and territory child protection bodies to share information (Cheng, 2019).

Information sharing among child protection and policing bodies is critical, but we believe the system needs to ensure information on children is gathered from multiple sources. This would include, for example, General Practice Physicians and medical centres, schools, child/day care centres, and counselling services. Particularly in relation to sexual offences against children, we know that disclosure rates are low, and bringing together information and observations from a range of sources is critical to ensure a complete picture.

One observation we have had working within this space is that identifying vulnerable and at-risk children is often analogous to a jigsaw puzzle, with various 'pieces' of information from numerous sources.

6-year-old 'Ben' has several interactions with various individuals. One of the workers at his after-school care notice that he is moody and isolating himself, he is also acting mildly inappropriately with one of the other children. His teacher notices that his grades have dropped, he seems distracted in class and his drawings have very dark overtones. His swimming teacher notices bruises on his lower back and upper thighs but thought they could be caused by something innocent. His babysitter notices that Ben seems to avoid being alone with his 13-year-old brother, who is actively aggressive towards him.

Bravehearts has previously discussed how an automated decision-making tool might be able to be utilised to bring together 'observations' from multiple sources and provide an assessment of risk that could then be reviewed. This is something that we believe could prove invaluable in providing those with or without mandatory reporting responsibilities with an opportunity to record observations or behaviours that on their own may not trigger a mandatory report.

This may also assist with concerns around mandatory reporting possibly resulting in over-reporting, and an influx of notifications to child protection bodies that require a review. Artificial Intelligence may provide an opportunity for necessitating different levels of human intervention when predetermined levels of risk are identified.

Predicting Risk

Accurate risk assessment is crucial in making decisions about the vulnerability of a child and the possibility of harm. However, there is no fool-proof method of assessing risk and no single instrument or data source in and of itself should be used to make critical decisions that impact on the safety and protection of children. Predicting the future is difficult and errors can include false negatives (leaving a child in danger) or false positives (removing a child who would not have been harmed).

Understanding that automated decision-making tools are also susceptible to errors is important. However, the benefits of an automated decision-making approach in the child protection sector are if it does not replace human intervention or assessment, but instead supports and supplements the assessment of risk. An automated decision-making tool could provide confirmatory support for assessments of risk, reducing the possibility of erroneous or biased human assessment. In addition, it could assist in objectively identifying risk indictors and correlations from diverse and multiple sources.

Openness, Transparency and Accuracy

As an NGO working in this space, a broad concern around current practices is the need for openness and transparency. This would also be an important consideration for the development of an automated decision-making tool and how the technology may be understood and 'trusted' by the sector.

We would suggest that the introduction of any technological tool be accompanied by a plain language explanatory, which provides an understanding of the 'inner workings' of predictive modelling/algorithms (inputs, outputs, weightings).

3. ROUNDTABLE DISCUSSION

Edited by Dr Lyndal Sleep

The presentations were followed by a round table discussion, where attendees asked questions and commented on the presentations. Attendees included the presenters, as well as researchers, and workers in child protection and the child and family sector from both government and community agencies. Attendees also included advocates of parents and children who have come into contact with child protection agencies. This section summarises the transcript of the discussion, including questions asked, responses to inquiries, and comments. The discussion is organised according to key thematic subheadings, providing scaffolding for the excerpts from the transcript. The excerpts have been lightly edited for readability, however the intention is to keep them as original as possible to retain the authentic voice of the participants.

In general, to protect the participants' confidentiality, the speakers are not identified in the transcript unless they have explicitly requested to be. Some speakers may be identifiable from the context of the conversation, however these were the presenters in the session and have given approval for this account of proceedings.

On monitoring/ evaluation of tools and unintended consequences

"What are some effective ways to ensure ongoing monitoring of the tools in terms of their application from a systemic level to make sure that they are contemporary and being used consistently?"

Quality assurance

"Post-deployment quality assurance must involve building a kind of back-end infrastructure around the use of predictive risk models that look at each of the features and test them before they go into the model. Then you have to follow decisions and do a quasi-experimental almost impact evaluation. CSDA has quarterly management reports of what people are doing in response to the risk scores. There's both a technical deployment, which requires a system that makes sure the tools are not broken because it's very difficult from the content to see if it is broken. This is important because a social worker will take that score as given, and you do not know if it is pulling the data in properly, so we have built a lot of quality assurance around that. We have software engineers and so on, and so we deploy a bunch of backend quality assurance that works to look for breaks and features breaks and scores. COVID has been a big challenge for us, so we've had to do substantial work with the agencies with COVID changing referral patterns, so we have done a lot of work in that."

Lack of transparency, the need for better communication and unintended consequences

"I (Rhema) would add to that some of the work we are doing is looking at where things are going wrong and how people are being negatively affected. Also, we are finding that something that should be happening that is not happening is greater communication by different government agencies from different countries about what kinds of lessons can be learned from systems that go wrong. Particularly when it comes to child welfare, there is limited transparency. For example, when a pilot ends and a local authority decides not to proceed with the system, there is not a lot of information sharing going on about what happened or why that decision was made."

"There is little research that considers unintended consequences - some of the negative impacts that these systems can cause. For example, recently the Data Justice Lab brought together civil society organisations who were concerned about the move to digital welfare systems and data systems in the UK. We identified substantial knowledge within communities about the kind of data that is being collected and shared. This can deter some people from seeking help when they need it because they fear negative outcomes due to the collection of certain kinds of data. However, they might not be fully aware of how their sensitive data is being shared. This is a major problem. Often, we do not know who has access to data, how it is being shared and combined, as well as the long-term implications of this. More work needs to be done focusing on unintended consequences and also in learning from where things are going wrong."

Community engagement

"About community engagement, what points of differences are there between members of the public and what we are seeing in academia, government or in the practice space? What is the capacity for further engagement of children and young people?"

"CSDA has been chosen as a project partner for UNICEF. They are interested in developing guidelines for using these tools with children to get better input from children and young people themselves. Hopefully, with UNICEF, we'll be putting out some reports along these lines shortly (if they have not got a draft out already). Also, sometimes academics and data activists don't necessarily grapple with the same challenges that people subject to the tools and the frontline workers do. For example, I did a lot of community engagement about data sharing when I was working in a shelter. One client of the shelter said "I'm bipolar and sometimes during my mania I'll be on the street yelling" and he said "if the police could only share data with my key mental health workers the police would know that I was going through an episode, and I wasn't, say, on Meths or something. Which is often what police always assume when I'm having periods of mania". His concern was that data was not being shared, rather than a concern around data sharing. I have seen as much concern in my community engagement about data not being shared, as data being shared when I speak to front line people who are in the community and who might be subject to these tools. When I talk to data activists, I never hear about concerns about data not being shared, I only hear about their concerns about data being shared – there is kind of an asymmetry."

About the child and family services context

"Whose needs does the data meet in child and family services? Does it also meet the needs of families and parents, including parents who are the third and fourth generation in the child protection system?"

"I wanted to comment about Philip's work on the structured decision-making tool currently being used in child and family services in Queensland - about the context in which the data is used. For example, in the Queensland courts, the child protection practitioner will take the information to court about why the child is being removed. All the way through assessment, they are collecting data on that parent and their circumstances to support their allegations as to why that child needs to be removed. I am not questioning the validity of that process. I am just questioning who is supporting the parent at that time to understand the situation that they are in.

Parents who have had previous involvement in the child protection system, for example, who are third and fourth generation in the child protection system, very strongly say, "it ends with me". However, the challenges that those parents face are significant. If they have previous involvement with the child protection system, that is seen as a risk. It is not seen as a reason to go in and do early intervention or provide support. This makes them a possible target for child removal in the current system, regardless of their actual family context."

Need to look at how poverty impacts families and communities who come under the attention of child protection in Queensland.

"Does increasing numbers of children coming into care mean that abuse and neglect are increasing, or is it indicative of the lack of responsive, flexible services to families who do not have social or familial support? My data says that parents who have children removed generally experienced mental health issues, domestic violence, and/or alcohol and abuse issues. Most have experienced trauma that has not been dealt with in the first instance.

Shame and hiding are also issues, especially in Aboriginal and Torres Strait Islander communities. If you mention child protection in any environment, from what the parents have told me, that is a good reason to hide your children because of the impact of generational trauma from stolen generation policies.

It is important to point out that parents do reach out for help but may do so clumsily because they don't know where to go or how to frame their issue. For many parents, child protection means removal. I feel that transparency with families is essential in the ADM space: explaining why you are collecting that data, giving it some context. It is also essential to understand that families and parents who have come into contact with child protection authorities are suspicious. A common characteristic of those families is that they have reached out at some point, and generally they were not responded to appropriately.

Transparency in ADM is essential. Parents want to know why you are asking those questions; they want to know their context. They are suspicious of systems that have been involved in their families and they are very aware that you can take their children. I would support the use of shared tools, providing the parents and families with context and explanations.

Families and practitioners working beside each other supports determination, supports social justice principles. Supporting social justice principles also provides an opportunity for psycho-social education for parents about the risks and looks at opportunities to explore solutions together. It also invites families into that space where they have the opportunity to identify vulnerabilities with the parent and work with them to access help, avoiding punitive ways of working on vulnerable families.

Is it a reason to intervene or is it a reason to provide support? Parents have said surveillance is support. Data, what is it good for and why are we using it? I feel very passionate about utilising the data finding tools to do an analysis of the systems that families are experiencing. I think that the reasons for poverty and the reasons I mentioned before for children coming into care could be addressed. We should be looking at ourselves as a system and how responsive we are being, because those child protection numbers are going up and they're not coming down."

Forensic investigation risk assessment models versus models which support the parents to care for their children better

"Child protection sits within the context of having an almost forensic investigation model which focuses on removing children rather than helping parents look after their children better, which is why the numbers of children in care are increasing. The threshold for removing children was higher in the past than it is today. We need more supportive services so that children can grow up at home safely, but this is expensive. We have gotten to a point where we can harm more and more children through a punitive child protection system."

A human rights approach to ADM in child protection

"One of the challenges to enforcing human rights-based decision-making is that people do not understand when they are making decisions, and this is the case in child protection systems too. Also, it is fundamental that people using the system, including parents and children who come into contact with the child protection system, play a role in the development of ADM in the sector - but this often does not occur. It is a significant challenge to determine how to make the best use of ADM when there are fundamental problems around the design of the current child protection system, definitional and conceptual. We need to avoid perpetuating these problems and have an opportunity to move forward solving them. The new Queensland Human Rights Legislation Act also states that Aboriginal and Torres Strait Islander children have a right to kinship care."

"Victoria has developed a new risk assessment framework. They are currently trying to have a social justice approach to child protection, which supports families to care for children. To not focus on child abuse and neglect but to locate families who need more support with their parenting, working with families to do this—currently redesigning training to focus more on supporting families with broader issues, including health navigator roles which support mental health, disability etc. The aim is to think differently about how we look at our professional judgement."

Beyond prescriptive risk assessment in child protection ADM

"Can ADM be used to identify supports for families who have come to the attention of child protection, that can reduce their risk scores?"

"Machine learning tools focus on three types of features: 1. features that predict risk, 2. features that are mutable but may not be very predictive, 3. features that are causative. When building tools for mutable factors, there is a different type of tool, sometimes called prescriptive analysis, in which evidence can suggest a program that could help the family. We are working on this in homelessness."

"When working with families in child protection, social workers using positive reinforcement models do not find current prescriptive risk assessment tools useful."

"It will be helpful to work out how to use unstructured data like case notes to give context in ADM systems. Also, it would be helpful to make use of metadata, which can take policy decisions into account in ADM, providing a broader policy context for decision-making."

4. BIOGRAPHIES

Dr Philip Gillingham practiced social work, mainly in child protection, at various levels for 8 years in England and 8 years in Victoria. He moved to an academic position at Deakin University in 2004 and completed his PhD at the University of Melbourne in 2009, *The use of assessment tools in child protection: an ethnomethodological study*. In 2011, he moved to the University of Queensland, where he is now an ARC Future Fellow and Associate Professor in social work.

Prof Paul Henman is Professor of Digital Sociology and Social Policy at the University of Queensland, and a Chief Investigator of the ARC Centre of Excellence for Automated Decision-Making and Society. His main research interest is the nexus between social policy, administration and digital information technologies.

Dr Joanna Redden is Assistant Professor, Faculty of Information and Media Studies at Western University, Ontario, Canada. She is also Co-director of the Data Justice Lab (Cardiff University) and co-author of *Compromised Data: From Social Media to Big Data and The Mediation of Poverty* (2015).

Carol Ronken is Director of Research at Bravehearts, Australia and Visiting Fellow, School of Justice, Faculty of Law, QUT.

Dr Lyndal Sleep is a Research Fellow at the ARC Centre of Excellence for Automated Decision-Making and Society's University of Queensland node. Her research focuses on social security decision-making, technology and women with the aim of enhancing safety, wellbeing and life chances for women in situations of intersectional disadvantage.

Prof Rhema Vaithianathan is a Professor of Social Data Analytics at the Institute for Social Science Research at The University of Queensland (UQ). She is also Director of the Centre for Social Data Analytics (CSDA) which is based at Auckland University of Technology, New Zealand (where she is a Professor of Economics) and has a second site at UQ. CSDA research projects are focused on leveraging data science to improve decision-making and outcomes and reduce disparities for vulnerable populations.

5. REFERENCES

Allegheny County Department of Human Services (2019). *Impact Evaluation Summary of the Allegheny Family Screening Tool.* https://www.alleghenycountyanalytics.us/wp-content/uploads/2019/05/Impact-Evaluation-Summary-from-16-ACDHS-26_PredictiveRisk_Package_050119_FINAL-5.pdf

Binns, R. (2017). Algorithmic Accountability and Public Reason. *Philosophy and Technology* 31, 1–14.

Amrit, C., Paauw, T., Aly, R., & Lavric, M. (2017). Identifying child abuse through text mining and machine learning. *Expert Systems with Applications* 88, 402–418.

Cheng, A. (2019). New AI platform for child protection agencies. Government News. https://www.governmentnews.com.au/new-ai-platform-for-child-protection-agencies/

Children's Data Network and California Department of Social Services (2020). *Assessing Children's Risk Using Administrative Records. A Proof of Concept Predictive Risk Modelling (PRM) Project.*

https://www.datanetwork.org/research/assessing-childrens-risk-using-administrative-records-a-proof-of-concept-predictive-risk-modeling-prm-project/

Chouldechova, A., Putnam-Hornstein, A., Benavides-Prado, D., Fialko, O. & Vaithianathan, R. (2018). A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. *Proceedings of Machine Learning Research* 81, 1-15.

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*, available: https://whatworks-csc.org.uk/wp-content/uploads/WWCSC_machine_learning_in_childrens_services_does_it_work_Sep_2020_Accessible.pdf

Courtland, R. (2018) Bias detectives: the researchers striving to make algorithms fair, 20 June, *Nature. https://www.nature.com/articles/d41586-018-05469-3*

Dare, T. (2013). Predictive Risk Modelling and Child Maltreatment: An Ethical Review. http://www.msd.govt.nz/documents/about-msd-and-our-work/publications-resources/research/predictive-modelling/00-predicitve-risk-modellingand-child-maltreatment-an-ethical-review.pdf

Dencik, L., A. Hintz, Z. Carey, et al. (2015). *Managing 'Threats': Uses of Social Media for Policing Domestic Extremism and Disorder in the UK.* Cardiff School of Journalism, Media and Cultural Studies, October. Accessed November 7, 2015. http://www.dcssproject.net/files/2015/12/Managing-Threats-Project-Report.pdf.

Devlieghere, J., Gillingham, P. & Roose, R. (forthcoming, 2021) *Dataism vs relationshipism: a social work perspective.* Special edition of *Nordic Social Work Research.*

Drake, B. & Jonson-Reid, M. (2018). *Administrative Data and Predictive Risk Modelling in Public Child Welfare: Ethical issues relating to California*. Available at http://www.caichildlaw.org/Misc/Ethical_Review_of_Predictive_Risk_Modeling.pdf

Eubanks, V. (2018). *Automating Inequality. How High Tech Tools Profile, Police and Punish the Poor*. New York: St. Martins Press.

Gillingham, P. (2009a). Practitioner Perspectives on the Family Risk Evaluation Tool: An aide to decision making or 'just another form to fill in'? *Developing Practice: the Child, Family and Youth Work Journal*, 23, 4, 46-55.

Gillingham, P. (2009b). *The use of assessment tools in child protection: an ethnomethodological study.* University of Melbourne. Published online at: http://repository.unimelb.edu.au/10187/4337

Gillingham, P. & Humphreys, C. (2010). Child protection practitioners and decision making tools: observations and reflections from the frontline. *The British Journal of Social Work*, 40, 8, 2598-2616.

Gillingham, P. (2011). Decision making tools and the development of expertise in child protection practitioners: are we 'just breeding workers who are good at ticking boxes?' *Child and Family Social Work*, 16, 4, 412-421.

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, 1044-1058.

Gillingham, Philip. (2015a). Implementing Electronic Information Systems in Human Service Organizations: The Challenge of Categorization. *Practice* 27 (3): 163–175.

Gillingham, P. (2015b). Electronic information systems in human service organisations: the what, who, why and how of information. *The British Journal of Social Work*, 45, 5, 1598-1613.

Gillingham, P. (2017). Predictive risk modelling to prevent child maltreatment: insights and implications from Aoteaora/New Zealand. *Journal of Public Child Welfare*, 11, 2, 150-165.

Gillingham, P. & Graham, T. (2017). "Big data" in social welfare: The development of a critical perspective on social work's latest "electronic turn". *Australian Social Work Special Issue: e-Professionalism and the Ethical Use of Technology in Social Work*, 70, 2, 135-147.

Gillingham, P. (2019). Algorithms, social policy, social justice and social work: principles of algorithmic accountability. *Practice: social work in action*, 31, 4, 1-14.

Gillingham, P. (2019). Can predictive algorithms assist decision making in social work with children and families?. *Child Abuse Review.*

Gillingham, P. (2020). Algorithmically based decision support tools: skeptical thinking about the inclusion of previous involvement. *Practice: Social Work in Action*. https://doi.org/10.1080/09503153.2020.1749584

Gillingham, P. (2020a). The development of algorithmically-based decision-making systems in children's protective services: is administrative data good enough? *The British Journal of Social Work*, 50, 565–580.

de Haan, I., and Connolly, M. (2014) Another Pandora's Box? Some pros and cons of predictive risk modelling. *Children and Youth Services Review* 47(1), 86–91.

Eberhardt, Jennifer L. (2020). *Biased: Uncovering the hidden prejudice that shapes what we see, think, and do*. Penguin Books

Hollnagel, E. and D.D. Woods (2005). *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering.* Boca Raton, FL: CRC Press.

HM Govt (2018). Information sharing, Advice for practitioners providing safeguarding services to children, young people, parents and carers (July 2018). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/721581/Information_sharing_advice_practitioners_safeguarding_services.pdf

Hornby Zeller Associates (2018). Allegheny County Predictive Risk Modelling Tool Implementation: Process Evaluation. https://www.alleghenycountyanalytics.us/wp-content/uploads/2019/05/Process-Evaluation-from-16-ACDHS-26_PredictiveRisk_Package_050119_FINAL-4.pdf

Hurley, A. (2018) Can an Algorithm Tell When Kids Are in Danger? *New York Times Magazine*, January 2018. https://www.nytimes.com/2018/01/02/magazine/can-an-algorithm-tell-when-kids-are-in-danger.html

Jackson, K. (2018) Predictive Analytics in Child Welfare — Benefits and Challenges, *Social Work Today*, 18(2), p. 10. https://www.socialworktoday.com/archive/MA18p10.shtml

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.

Kelly, J (2015). Florida's New Predictive Risk Tool Likely to Drive Down Juvenile Incarceration (23/6/2015) *The Imprint* https://imprintnews.org/featured/floridas-new-predictive-risk-tool-likely-to-drive-down-juvenile-incarceration/10505

Kitchin, R. (2014). *The Data Revolution.* London: Sage.

Kitchin, R. (2015). *Data-Driven, Networked Urbanism, The Programmable City* Working Paper, 14. August 10.

Leslie, D., Holmes, L., Hitrova, C. & Ott, E. (2020). *Ethics review of machine learning in children's social care. What works in children's social care?* Rees Institute, University of Oxford/The Alan Turing Institute.

Munroe, E. (2010). The Munro Review of Child Protection. Part One: A Systems Analysis. Department for Education, Oct. 1. Accessed December 10, 2016. URL: https://www.gov.uk/government/publications/munro-review-of-child-protection-part-1-a-systems-analysis.

Martin, E. K., & Silverstone, P. H. (2013). How much child sexual abuse is "below the surface," and can we help adults identify it early? *Frontiers in Psychiatry*, 4, 1-10.

Mills, R., Kisely, S., Alati, R., Strathearn, L., & Najman, J. (2016). Self-reported and agency-notified child sexual abuse in a population-based birth cohort. *Journal of Psychiatric Research*, 74, 87-93

Pedersen, J.S. (2019). Data-driven management in practice in the digital welfare state. Chapter 10 in J.S. Pedersen and A. Wilkinson, *Big Data: Promises, Application and Pitfalls.* Edwin Egar: England, pp. 200-223.

Predictive Risk Modelling: Practical Considerations (2020) https://www.datanetwork.org/wp-content/uploads/PRM_Final.pdf

Price-Robertson, R., Bromfield, L., & Vassallo, S. (2010). Prevalence matters: Estimating the extent of child maltreatment in Australia. *Developing Practice: The Child, Youth and Family Work Journal*, 26, 12-20.

Redden, J., Dencik, L. and Warne, H. (2020). Datafied child welfare services: unpacking politics, economics and power, *Policy Studies*, DOI: 10.1080/01442872.2020.1724928

Royal Commission into Institutional Responses to Child Sexual Abuse (2017). Final report, Volume 4: Identifying and disclosing child sexual abuse. Commonwealth of Australia.

Salganik, M.J. (2018). *Bit by Bit: Social Research in the Digital Age.* Princeton University Press: Princeton, NJ.

Tarzon, C. & Quadara, A. (2012). The Nature and Extent of Sexual Assault in Australia. (Resource Sheet). Melbourne [Vic]: Australian Centre for the Study of Sexual Assault.

Taylor, Brian J. (2017). Heuristics in professional judgement: A psycho-social rationality model. *British Journal of Social Work*, 47(3), 1043-1060.

USACM (2017). *Statement on Algorithmic Transparency and Accountability.* US Association for Computing Machinery Public Policy Council (USACM). https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf

Vaithianathan, R., Maloney, T., Putnam-Hornstein, E. and Jiang, N. (2013). Children in the public benefit system at risk of maltreatment: Identification via predictive modelling. *American Journal of Preventative Medicine* 45(3), 354–359.

Vaithianathan, R., E. Putnam-Hornstein, N. Jiang, P. N. and, and Maloney. T. (2017). Developing Predictive Models to Support Child Maltreatment Hotline Screening Decisions: Allegheny County Methodology and Implementation. Accessed January 8, 2020. https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx

Vaithianathan, R., Putnam-Hornstein, E., Chouldechova, A., Benavides-Prado, D., & Berger, R. (2020). Hospital injury encounters of children identified by a predictive risk model for screening child maltreatment referrals: evidence from the Allegheny Family Screening Tool. *JAMA Pediatrics*, e202770-e202770

White, S., Wastell, D., Broadhurst, K., & Hall, C. (2010). When Policy O'erleaps Itself: The 'Tragic Tale' of the Integrated Children's System. *Critical Social Policy* 30(3): 405–429.

Useful links

Centre for Social Data Analytics Website (AUT): https://csda.aut.ac.nz/

Vaithianathan, R., Putnam-Hornstein, E., Chouldechova, A., Benavides-Prado, D., & Berger, R. (2020). Hospital injury encounters of children identified by a predictive risk model for screening child maltreatment referrals: evidence from the Allegheny Family Screening Tool. *JAMA pediatrics*, e202770-e202770, https://jamanetwork.com/journals/jamapediatrics/article-abstract/2769075

Plain English reports: Predictive Risk Modeling for Child Protection (December 2019) https://csda.aut.ac.nz/__data/assets/pdf_file/0013/432220/Predictive-Risk-Modeling-FS-1.pdf

Advanced Analytics for Child Welfare: Spotlight on Supervision Support (June 2020) with Children's Data Network (USC) and Mathematica (US).
https://csda.aut.ac.nz/__data/assets/pdf_file/0003/432219/Analytics-Child-Welfare-FS.pdf

Bias detectives: the researchers striving to make algorithms fair, June 2019, Nature News, https://www.nature.com/articles/d41586-018-05469-3

Predictive Analytics in Child Welfare — Benefits and Challenges, *Social Work Today*, March/April 2018, https://www.socialworktoday.com/archive/MA18p10.shtml

Can an Algorithm Tell When Kids Are in Danger? *New York Times Magazine*, January 2018, https://www.nytimes.com/2018/01/02/magazine/can-an-algorithm-tell-when-kids-are-in-danger.html admscentre.org.au

Australian Government
Australian Research Council

This Centre is funded by the Australian Government through the Australian Research Council

ADM +S