Researching AI Applications in Child Welfare

Joanna Redden	Shion Guha
Western University	University of Toronto
jredden2@uwo.ca	shion.guha@utoronto.ca

Overview

1) Context and relevant research
2) Cancelled system project findings - JR
3) Case study findings - SG
4) Closing

MACHINE LEARNING

Can machine learning save children at risk?

The idea that machine learning can improve children's social care is attractive, but fraught with challenges – the veil of secrecy around predictive analytics in public services must be lifted, say **Michael Sanders** and **Vicky Clayton**

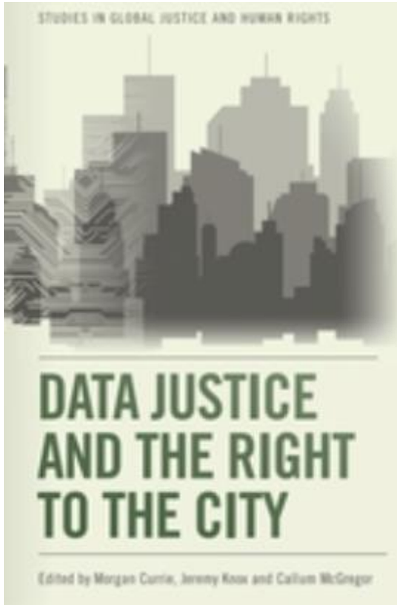
What Works *for* Children's Social Care

MACHINE LEARNING IN CHILDREN'S SERVICES

SUMMARY REPORT

Vicky Clayton
Michael Sanders
Eva Schoenwald
Lee Surkis
Daniel Gibbons

SEPTEMBER 2020



Chapter 3

DATAFIED CHILD WELFARE SERVICES AS SITES OF STRUGGLE

Joanna Redden, Jessica Brand, Ina Sander and Harry Warne

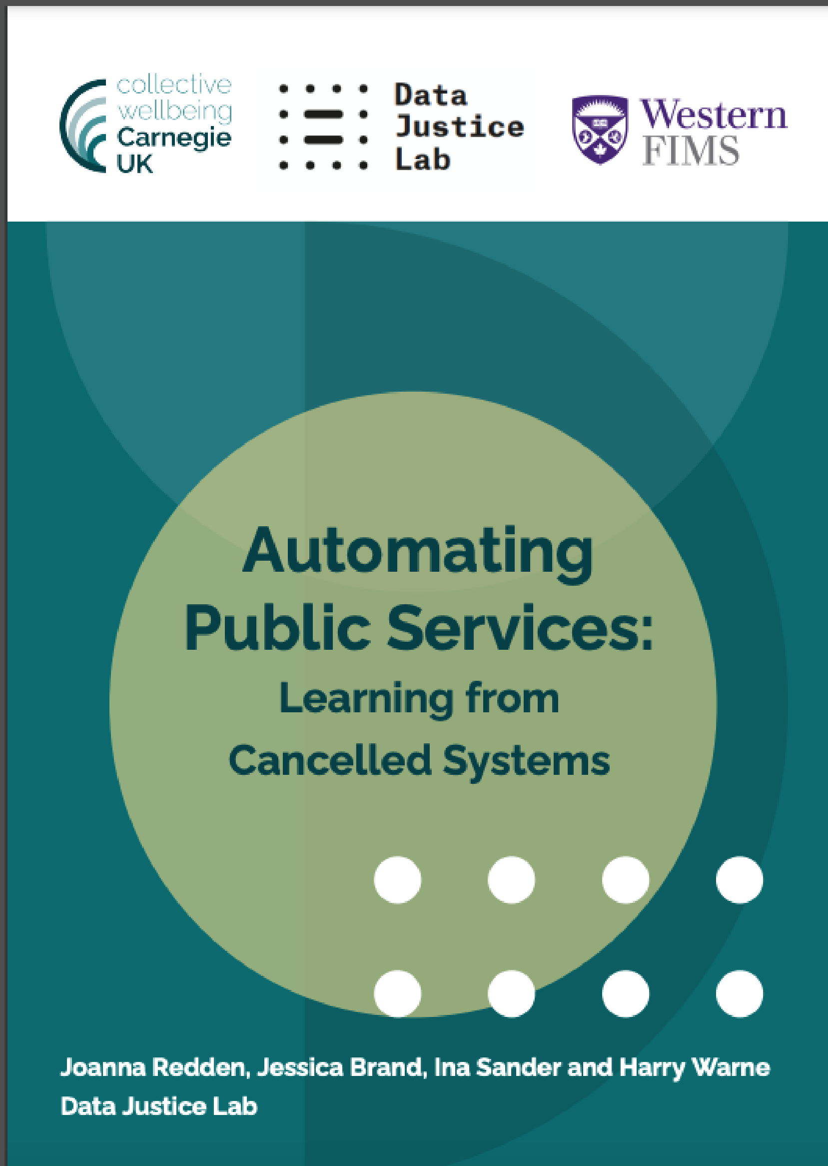
POLICY STUDIES
2020, VOL. 41, NO. 5, 507–526
https://doi.org/10.1080/01442872.2020.1724928

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Datafied child welfare services: unpacking politics, economics and power

Joanna Redden, Lina Dencik and Harry Warne



STARLING CENTRE

JUST TECHNOLOGIES
JUST SOCIETIES

Factors influencing decisions to cancel

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

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).

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).

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)

Case Studies

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).

Political economic context matters

- UK Austerity context significant
- Political economy: “troubled families” policy as driving force
- Governmentalities, legalities and questions about rights
- The public sector as datafied marketplace
- Differing levels of transparency for ‘knowledge’ outputs
- Materialities, infrastructures and practices: the importance of historical context
- Organizations, communities, places and subjectivities: attending to agency and resistance

Redden J., Dencik, L. and Warne H. (2020) “Datafied child welfare services: Politics, economics and power,” *Policy Studies*, Special Issue “Political Economy of Digital Data,” Prainsack, B. (ed.).

Recommendations

- 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

When Prediction Becomes Policy: What Do Risk Models Actually Learn?

Why we cannot predict “risk” without modelling the intervention.

Shion Guha

Assistant Professor
Faculty of Information
Department of Computer Science
Director, Human-Centered Data Science Lab
University of Toronto

RESEARCH-ARTICLE

A Framework of High-Stakes Algorithmic Decision-Making for the Public Sector Developed through a Case Study of Child-Welfare

Published: 18 October 2021

Citation in BibTeX format

DEVANSH SAXENA, Marquette University, Milwaukee, WI, United States

KARLA A. BADILLO-URQUIOLA, University of Central Florida, Orlando, FL, United States

PAMELA J. WISNIEWSKI, University of Central Florida, Orlando, F
United States

SHION GUHA, University of Toronto, Toronto, ON, Canada

RESEARCH-ARTICLE

Rethinking "Risk" in Algorithmic Systems Through A Computational Narrative Analysis of Casenotes in Child-Welfare

DEVANSH SAXENA, Marquette University, Milwaukee, WI, United States

ERINA SEH-YOUNG MOON, University of Toronto, Toronto, ON, Canada

ARYAN CHAURASIA, University of Toronto, Toronto, ON, Canada

YIXIN GUAN, University of Toronto, Toronto, ON, Canada

SHION GUHA, University of Toronto, Toronto, ON, Canada

Open Access Support provided by:

University of Toronto

Marquette University

Published: 19 April 2023

Citation in BibTeX format

CHI '23: CHI Conference on Human Factors in Computing Systems
April 23 - 28, 2023
Hamburg, Germany

Very recently: Translating to Ontario, Canada (CHI'26?)

Concerns about risk assessment standards and tools

Consultation participants raised concerns about bias in the tools and standards used to assess risk to children. Although they seem neutral, we heard that risk assessment standards and tools may lean towards more positive outcomes for White people.

Social work researchers argue that risk assessment tools in Ontario are biased and perpetuate racism because they do not account for structural inequalities, such as racial discrimination, that may affect a child's well-being. Parents may be blamed for these external factors, even though they are largely out of their control. We heard that relying on these tools, coupled with worker bias – which may be conscious or unconscious – may contribute to assumptions about racialized children and families being "inherently wrong or deficient." This can lead to incorrect assumptions about the level of risk children are exposed to.

We also heard concerns about risk assessment standards that relate to poverty – for example, the number of children allowed per bedroom. Poverty in racialized and Indigenous families may be seen as a sign of neglect, providing a basis for a child welfare agency to become involved. We heard that these standards can affect what is seen as acceptable in a home and contribute to CAS decisions to intervene.

It is unclear to what extent child welfare risk assessment standards and tools reflect real risk to children in all cases, or arise from White, Western, Christian middle-class norms. When standards and tools are not based on objective factors, but on the cultural norms of the dominant group, they may contribute to racial profiling.

Concerns about biased decision-making

Concerns were also raised both about the perceived bias of authorities or individuals that refer to CASs, and perceived bias in decision-making practices when child welfare workers and authorities become involved with families. Participants said that child welfare workers, many of whom are White, may be more likely to construe family situations or the actions of Indigenous or racialized people as "risky."

The Ontario Federation of Indigenous Friendship Centres (OFIFC) identified that Indigenous families experience "intense scrutiny of [their] ways of life" (for more information, see the full OHRC report, *Under suspicion: Research and consultation report on racial profiling in Ontario*). We repeatedly heard that non-Indigenous child

welfare workers often do not understand the nature or structure of Indigenous families and cultural differences in how families live. For example, they only see that children are not being raised by their parents or are living in what they think are over-crowded conditions. In another example, Indigenous youth told us that they are sometimes put into care because they miss a lot of school due to practicing their traditions and taking part in ceremonies.

Social work researchers talked about some of the factors that may contribute to the over-scrutiny of Black parents, and the tendency to view Black parents as risks to their children and in need of intervention by CASs. For example, researchers note tha[...] TO TOP [...]uthorities commonly view Black parents as "aggressive" and "crazy" when they

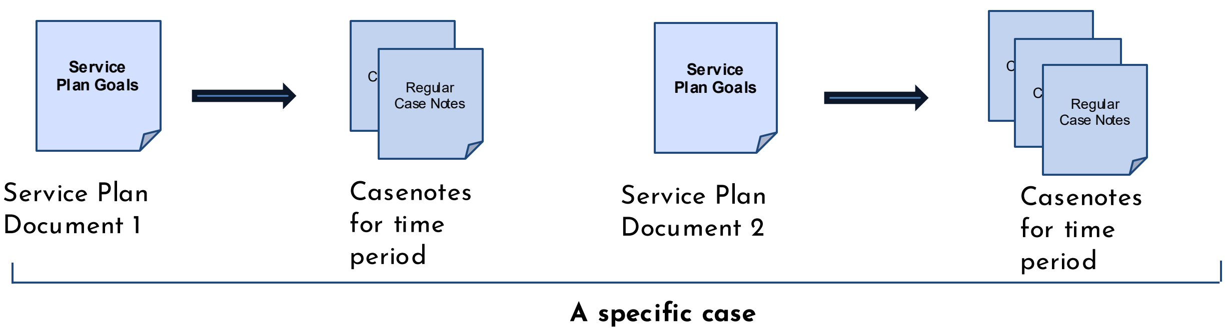
(L-R) Matthew Tamura, Shan (Angelina) Zhai, Professor Shion Guha (Faculty of Information, University of Toronto) worked on the Children's Aid Society MITACS project

CHILDREN'S AID SOCIETY of TORONTO | LA SOCIÉTÉ DE L'AIDE à L'ENFANCE DE TORONTO

Supporting CW workers in Canada

(Under review)

Child welfare agencies want to ensure workers are providing effective services for families and reduce long-term case backlog



Approach:

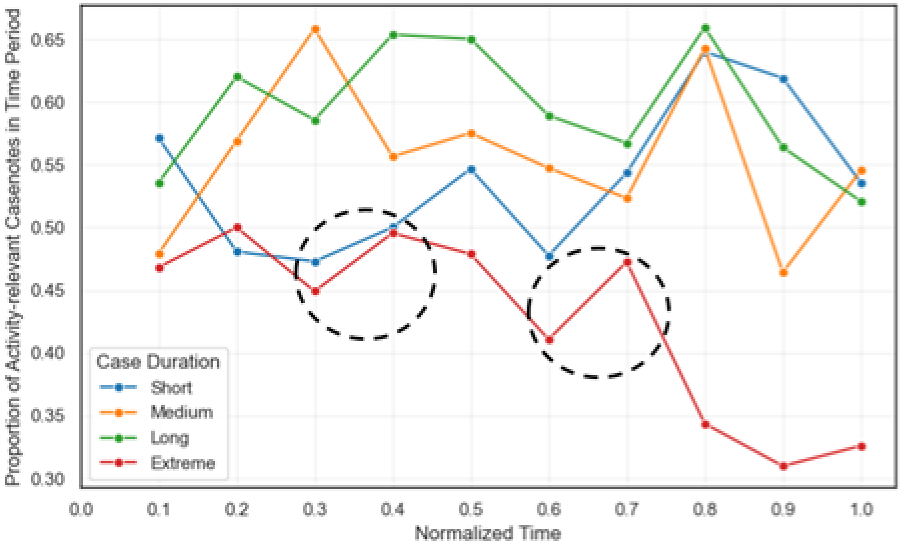
- Identify regular casenotes relevant to goals using LocalLLM (Llama 3.1) and manual labeling with child welfare practitioners
- Trace thematic trajectory of Regular casenotes using BERTopic

Please read, reuse and cite: https://tinyurl.com/simons-ucberkeley | shion.guha@utoronto.ca | https://www.hcds-uoft.ca

Supporting CW workers in Canada

(Under review)

- Divided casenotes by case duration
- Extremely long term cases make fewer references to Service Plan Goals compared to shorter term cases
- Black circles denote when new Service Plans are drawn up for a case in Extreme cases



Supporting CW workers in Canada

(Under review)

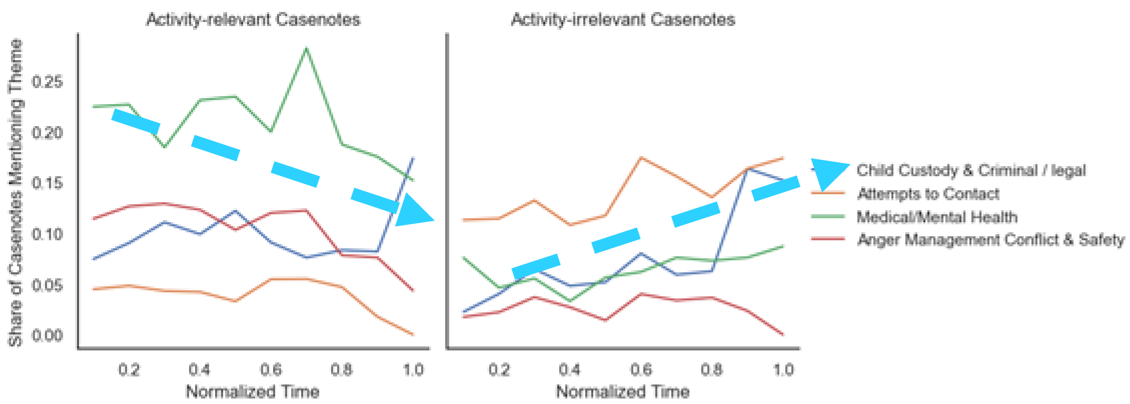
Manual labeling vs. LLM

- LLMs struggle at identifying Service Plan goal relevant Regular Casenotes as cases become more complex and uncertainty increases

Case Duration	Agreement	Cohen's κ	FPR	FNR
Short	0.804	0.604 [0.535, 0.670]	0.265	0.135
Medium	0.782	0.550 [0.493, 0.602]	0.324	0.135
Long	0.727	0.402 [0.352, 0.447]	0.470	0.146
Extreme	0.730	0.470 [0.440, 0.500]	0.382	0.134

Tracking thematic Service Plan goal relevance using BERTopic

- Caseworkers are working on new child welfare concerns that emerge in Extremely long cases that are not outlined Service Plans (right plot)



Human-Centered Data Science Lab University of Toronto

MAIN TAKEAWAYS

Risk models often learn the intervention system, not latent harm.

- Administrative outcomes are often **system contact labels**, produced by human discretion and bureaucracy.
- Narrative records surface **intervention work, constraints and power** which predictive risk models typically erase
- The meaning of risk is **temporal and process dependant**; static risk scoring collapses actual human dynamics into a number.
- LLMs do not solve this in long, complex cases; **reliability declines as uncertainty and case entropy rise**.
- When prediction ignores intervention, it becomes punishment.

Human-Centered Data Science Lab
University of Toronto