Exploring the role of Generative AI models for research in the public sector

Erina Seh-Young Moon
University of Toronto
Toronto, Canada
erina.moon@mail.utoronto.ca

Kelly McConvey
University of Toronto
Toronto, Canada
kelly.mcconvey@mail.utoronto.ca

Shion Guha
University of Toronto
Toronto, Canada
shion.guha@utoronto.ca

ABSTRACT

The rise of generative artificial intelligence (GenAI) opens up new affordances in human-computer interaction (HCI) research on public sector sociotechnical systems. We argue that there remain opportunities for using AI to generate synthetic data when theoretically assessing the fairness of decisions rendered by AI systems in public sector domains. However, there remain critical ethical challenges to adopting generative AI for public sector research if synthetic data or personae substitute stakeholder voices. In this paper, we draw on past research from child welfare, higher education, and homelessness systems to identify opportunities and challenges in adopting generative AI for research in the public sector.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; • **Applied computing** → **Computing in government**.

KEYWORDS

artificial intelligence, generative AI, homelessness, child welfare, higher education, public sector

ACM Reference Format:
Erina Seh-Young Moon, Kelly McConvey, and Shion Guha. 2024. Exploring the role of Generative AI models for research in the public sector. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24).* ACM, New York, NY, USA, 3 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Public sector services and institutions in the US and Canada have long used information communication technologies (ICTs) to increase public service delivery efficiencies, minimize repeated information collection, reduce bureaucratic overhead, and promote consistent and evidence-based decision-making processes [4, 5]. Following the recent rise of generative AI models, critical questions are being raised about how public sector workers can and should interact with powerful nascent technologies such as Large Language Models (LLMs). In fact, in Canada, the federal government recently introduced guidelines on how federal institutions can use these tools in public work in recognition of their powerful potential and limitations [3]. Because the nature of public sector services and work often involves critical matters impacting citizens and dealing with highly sensitive personal information, public agencies in the US and Canada emphasize the need to exercise caution with what information is input into generative AI tools and how the tools' outputs are trusted or used.

Prior ethnographic work by HCI scholars in the public sector highlights the dynamic and complex interactions between technology use and human decision-making. Workers are subject to varying bureaucratic protocols and legislative and organizational constraints, which impact how workers use technologies. Often, researchers have found that in the face of the above-mentioned constraints, AI tools fuse onto existing work practices, become misappropriated for uses beyond their initial intended use, and inhibit discretionary human decision-making, leading to poor decision-making [7, 12]. Additionally, stakeholders who use and are impacted by decisions rendered by AI systems raise critical concerns on the appropriateness of using black-box systems to alter an individual's life significantly [8, 15]. Given the powerful capabilities of generative AI tools and dynamic organizational or legislative constraints that significantly influence how AI tools are used in the public sector, it becomes critical to identify domain-specific opportunities and challenges when adopting generative AI to assist and regiment public sector work. In the following paragraphs, draw on past research from prior work studying child welfare, higher education, and homelessness systems to identify opportunities and challenges in adopting generative AI for research in the public sector.

2 CHALLENGES AND LIMITATIONS OF GENERATIVE AI FOR HOMELESSNESS, CHILD WELFARE, AND HIGHER EDUCATION SYSTEMS RESEARCH

Students of public Higher Education Institutions (HEIs), individuals experiencing homelessness and families involved with child welfare systems represent vulnerable communities. At the same time, these public sector domains face significant resource constraints. For example, child welfare systems face a shortage of experienced caseworkers and good foster homes [11], and cities experiencing high levels of homelessness do not have enough emergency shelter and affordable permanent housing spaces for their clients [8, 10]. In response to the resource-constrained landscape of child welfare, higher education, and homeless systems, public sector agencies and institutions are increasingly adopting AI tools to identify and prioritize high-risk populations and allocate services to them [9, 10, 14].

CHI '24, May 11–16, 2024, Honolulu, Hawai'i

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

In the face of the growing applications of generative AI, our prior studies examining caseworker work practices in child welfare systems highlight significant hurdles and ethical concerns of adopting these new powerful technologies [13, 14]. Our works previously examined narrative case notes written by front line staff in child welfare systems, which uncovered discretionary work carried out by caseworkers and complex power imbalances between child welfare system stakeholders (e.g., bio-parents, foster parents, children, caseworkers, judges, etc.). Through our work, we identified how caseworkers navigate the dual, conflicting role as advocates for families and enforcers of child welfare regulations. We also found caseworkers performed street-level discretionary work, such as helping clients with medical paperwork which is not officially documented as their job responsibility, to better serve their clients. Given the complex and dynamic interactions between multiple stakeholders in child welfare systems, questions arise about whether generative AI tools such as LLMs can begin to accurately represent the above mentioned tensions and dynamics that underpin child welfare cases.

Previous studies have found that LLMs can generate convincing and eloquent synthetic personae – i.e., producing human-like responses when prompted [6, 16]. However, LLM responses and personalities have also been found to be tied to the training data used [17]. If biological parents within child welfare systems are facing a range of unique challenges that require customized care from caseworkers [14], the practical utility of using generative AI such as LLMs to generate synthetic personae to express an individual's circumstances is low. Furthermore, serious ethical concerns arise if HCI researchers are to use synthetic personae in child welfare research instead of centering actually impacted stakeholder voices.

At the same time, we argue there remain opportunities for using generative AI to generate synthetic data when theoretically assessing the fairness of decision-rendered by AI systems in public sector domains. As noted above, many public sector agencies, including homeless systems and higher education, increasingly use algorithmic systems for decision-making. Our recent review of algorithmic systems designed for homelessness research [10] found current algorithmic models in homelessness focus on assessing an individual's risk of experiencing homelessness and conflate risk assessment with service delivery – i.e., high-risk individuals would be prioritized to receive services. We also found current algorithms failed to account for the resource-constrained landscape of homelessness and important fairness-related questions on how we measure risk, who should be allocated what type of service given a certain risk metric, and how these decisions will impact different fairness metrics were largely ignored in the design of algorithms for homelessness. LLMs could, therefore, offer affordances to generate synthetic data and run simulation studies to assess how different resource allocation decisions impact certain population groups and fairness metrics when operating within various resource-constrained public sector systems. Running randomized control trials, on the other hand, that allocate different support services can be costly, time-consuming, and not yield useful results. Utilizing generative AI tools could thus offer an alternative approach to studying how to design fair decision-making AI systems for public sector domains.

Similar to child welfare and homeless systems, public institutions of higher education have seen a significant increase in the deployment of educational data mining, learning analytics (which involves collecting data on a student's activity within the Learning Management System) [1], and the design of algorithms for decision-making, prediction, and personalization [2]. In our prior review of algorithmic decision-making in higher education [9], we demonstrated that model design has shifted away from rules-based systems of the 1990s and early 2000s towards neural networks and natural language processing. Consequently, the outcomes have become less explainable and interpretable, while requiring the use of greater quantities of individual student data, and subsequently increasingly raising privacy concerns. Despite this increase in both data and complexity, we found that these datasets and algorithmic systems largely lack context, treating student behaviours (such as learning management system activity), grades, and pathways as consistent across courses, programs, and institutions. Synthetic training data, such as that generated with LLMs, could potentially be used to develop student narratives that could be used to train more context-aware models without risking student privacy. However, we add a note of caution. The use of AI in higher education is already associated with exacerbating existing inequities, such as in the case of algorithmic Early Warning Systems [9]. And biases that exist within LLMs and their training data could reinforce these inequities. Just as with child welfare, serious ethical concerns arise if HCI researchers are to substitute LLM-generated synthetic training data in Higher Education research instead of centering student voices.

3 CONCLUSION

The rise of generative AI opens up new affordances in HCI research on public sector sociotechnical systems. Generative AI offers opportunities to generate realistic instances of synthetic data to run experimental simulation studies that would otherwise be costly and time-consuming. However, there remain critical ethical challenges to using generative AI for public sector research if synthetic data or personae substitute stakeholders' voices. We, therefore, need further exploration of the limitations and opportunities of using generative AI technologies for the public sector.

REFERENCES

[1] Nasir Abdul Jalil and Mikkay Wong Ei Leen. 2022. Learning Analytics in Higher Education: The Student Expectations of Learning Analytics. In *Proceedings of the 2021 5th International Conference on Education and E-Learning* (Virtual Event, Japan) *(ICEEL '21)*. Association for Computing Machinery, New York, NY, USA, 249–254. https://doi.org/10.1145/3502434.3502463

[2] Swati Bajpai and S. Mani. 2017. Big Data in Education and Learning Analytics. *TechnoLearn: An International Journal of Educational Technology* 7, 1and2 (2017), 45. https://doi.org/10.5958/2249-5223.2017.00005.5

[3] Canada. 2024. Guide on the use of generative AI. https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html

[4] Virginia Eubanks. 2018. *Automating inequality: How high-tech tools profile, police, and punish the poor.* St. Martin's Press.

[5] Sergio Fernandez and Hal G Rainey. 2017. Managing successful organizational change in the public sector. In *Debating public administration*. Routledge, 7–26.

[6] Perttu Hämäläinen, Mikke Tavast, and Anton Kunnari. 2023. Evaluating Large Language Models in Generating Synthetic HCI Research Data: a Case Study. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 433, 19 pages. https://doi.org/10.1145/3544548.3580688

[7] Naja Holten Møller, Irina Shklovski, and Thomas T Hildebrandt. 2020. Shifting concepts of value: Designing algorithmic decision-support systems for public services. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. 1–12.

[8] Tzu-Sheng Kuo, Hong Shen, Jisoo Geum, Nev Jones, Jason I. Hong, Haiyi Zhu, and Kenneth Holstein. 2023. Understanding Frontline Workers' and Unhoused Individuals' Perspectives on AI Used in Homeless Services. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3544548.3580882

[9] Kelly McConvey, Shion Guha, and Anastasia Kuzminykh. 2023. A Human-Centered Review of Algorithms in Decision-Making in Higher Education. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (, Hamburg, Germany,) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 223, 15 pages. https://doi.org/10.1145/3544548.3580658

[10] Erina Seh Young Moon and Shion Guha. 2024. A Human-Centered Review of Algorithms in Homelessness Research. (2024). https://doi.org/10.1145/3613904.3642392 arXiv:arXiv:2401.13247

[11] Devansh Saxena, Karla Badillo-Urquiola, Pamela Wisniewski, and Shion Guha. 2021. A Framework of High-Stakes Algorithmic Decision-Making for the Public Sector Developed through a Case Study of Child-Welfare. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021).

[12] Devansh Saxena, Karla Badillo-Urquiola, Pamela J Wisniewski, and Shion Guha. 2020. A Human-Centered Review of Algorithms used within the US Child Welfare System. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.

[13] Devansh Saxena, Seh Young Moon, Aryan Chaurasia, Yixin Guan, and Shion Guha. 2023. Rethinking "Risk" in Algorithmic Systems Through A Computational Narrative Analysis of Casenotes in Child Welfare *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, 19 pages. https://doi.org/10.1145/3544548.3581308

[14] Devansh Saxena, Seh Young Moon, Dahlia Shehata, and Shion Guha. 2022. Unpacking Invisible Work Practices, Constraints, and Latent Power Relationships in Child Welfare through Casenote Analysis *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 120, 22 pages. https://doi.org/10.1145/3491102.3517742

[15] Logan Stapleton, Min Hun Lee, Diana Qing, Marya Wright, Alexandra Chouldechova, Ken Holstein, Zhiwei Steven Wu, and Haiyi Zhu. 2022. Imagining new futures beyond predictive systems in child welfare: A qualitative study with impacted stakeholders. In *FAccT '22: 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul, Republic of Korea, June 21 - 24, 2022*. ACM, 1162–1177. https://doi.org/10.1145/3531146.3533177

[16] Mikke Tavast, Anton Kunnari, and Perttu Hämäläinen. 2022. Language Models Can Generate Human-Like Self-Reports of Emotion. In *27th International Conference on Intelligent User Interfaces (IUI '22 Companion)*. Association for Computing Machinery, New York, NY, USA, 69–72. https://doi.org/10.1145/3490100.3516464

[17] Fuzhao Xue, Yao Fu, Wangchunshu Zhou, Zangwei Zheng, and Yang You. 2023. To Repeat or Not To Repeat: Insights from Scaling LLM under Token-Crisis. arXiv:2305.13230 [cs.LG]