



Data Harms: The Evidence Against Education Data

Luci Pangrazio¹

Accepted: 6 March 2024
© Crown 2024

Keywords EdTech · Datafication · Bias · Digital privacy · Harms · Teacher labour

Introduction

Time and time again, we see the risks of collecting huge amounts of personal data about students. In the news, cyberattacks and data breaches at schools make headlines, and there is a body of academic research dedicated to critical studies of education data (Williamson 2019; Jarke and Breiter 2019; Selwyn et al. 2018). However, in schools and education departments, the role of data is rarely questioned—it provides granular insights into student learning, it can help predict performance and incentivise students, and it can be used to personalise learning and educational experiences.

Data gathering became crucial to new regimes of ‘accountability’ and ‘new public management’ that emerged in the late 1990s, and aligns with neoliberal approaches to education that rely on performance measures and top-down surveillance of teachers and students (Ball 2003; Connell 2013). However, critical studies of EdTech are beginning to provide evidence that education data can cause harm. Some emerging issues are due to the fact that personal data has become such a valuable commodity (Couldry and Mejias 2019; Zuboff 2019), while the composition of datasets and the design of algorithms present further problems (Noble 2018). It is an important area of inquiry given children and young people in school are subject to the implications of data often without knowing and typically with few opportunities to speak back and resist.

For now, the harms of education data can be grouped into at least four (often overlapping) categories, which can be summarised as bias, student alienation, increased teacher labour, and the erosion of children’s digital privacy. While there are other conceptual or theoretical issues associated with education data, the four categories discussed here are impacting people in schools *now* and should be cause for concern. The goal of this commentary is to raise education data as a social justice issue for key stakeholders working in schools, schooling systems, and education ministries.

✉ Luci Pangrazio
luci.pangrazio@deakin.edu.au

¹ Centre for Research for Educational Impact, Deakin University, Melbourne, Australia

Bias

Bias in data-driven systems is typically caused by prejudiced assumptions in algorithm development and/or the composition of the dataset that the algorithm is trained on. It has both statistical and social dimensions and can lead to inaccurate measurements and representations (Perrotta 2023). A growing body of research in education is documenting the harms from the bias in data-driven systems. A study by Dixon-Roman et al. (2020) into a common writing app Essay Helper showed that the algorithm is trained on data from testing regimes and Common Core Standards in the US that already privilege dominant (white) ways of writing and thinking, placing students from diverse language groups at a disadvantage. If schools make use of this app compulsory then prejudice becomes embedded in the education system.

Similar findings have been reported by Clutterbuck and colleagues (2021: 103) into an administrative EdTech platform in Queensland, Australia, called One School, in which students' indigenous languages could not be recorded 'alienating some students and their identities in schooling settings'. Lu and colleagues (2021: 2) reported biases are embedded in the EdTech platform ClassDojo, which 'disproportionately burden students experiencing marginalisation' due to their racial backgrounds, meaning they were more likely to be 'disciplined' than their white classmates.

These studies highlight the problematic privileges and representations embedded in some datafied systems, which can either intensify issues that already exist or create new ones. Importantly, these issues have real-world implications. Without accurate representation, it is difficult for schools and teachers to meet the specific needs and abilities of students from diverse backgrounds, which effects their educational outcomes and future opportunities.

Student Alienation

The second data harm is the *student alienation* caused by increased monitoring and surveillance through EdTech platforms and data. While the drive to monitor and improve student performance via data tracking is advertised as a benefit of many EdTech platforms, only recently has the impact on students been examined. In particular, the rise of student activity monitoring software (SAMS) has raised surveillance to a new level. SAMS are typically installed on a student's laptop by the school and enable teachers to view exactly which website a student is on at any time. Monitoring is in real-time and can be aggregated into class dashboards with simplified displays for teachers to check. Some SAMS enable teachers to turn off the Internet on student laptops or allow websites to be visited that might have been banned by the school. While SAMS enable greater control and oversight for teachers, students themselves report feeling powerless and despondent in relation to these technologies, which negatively impacts their feelings towards school, learning, and their teachers (Pangrazio et al. 2023a; Manolev et al. 2019).

SAMS is only one piece of software implemented by schools. Most schools also have Learning Management Systems (LMSs) that continuously collect

attendance data, as well as Learning Delivery Platforms (LDPs) like Canvas and Moodle that track visits to the site, activities completed, and assignment submissions. Given this level of surveillance and monitoring, it is not surprising that students might feel that teachers and parents are suspicious of them. The constant reinforcement of data encourages teachers to see students as a set of numbers or performance indicators rather than people. Trust is crucial to building relationships between staff and students and this may be jeopardised when students feel like they are under surveillance with teachers and parents watching every move.

Increased Teacher Labour

The third category concerns the *increased teacher labour* involved in making datafied systems work. Despite the promises of increased efficiencies and detailed insights into learning and schooling processes, a strand of critical EdTech research has in fact found that a lot of additional teacher labour is required to make these systems work. This additional labour comes in a range of forms, such as the need for specialist skills and time to enable data interoperability between different platforms (Pangrazio and Sefton-Green 2024; Pangrazio et al. 2023b), the different types of labour involved in operating platforms (Selwyn 2021), the different data roles or ‘classes’ that exist in schools (Selwyn et al. 2022), and the new divisions of labour that are emerging (Perrotta et al. 2021).

Not all these findings are necessarily negative, but they do indicate the expanding set of skills required, as well as the increased demands on resources and time. Across these studies, teachers who had skills and experience working with data often had a career advantage and were given positions of power and authority in schools. Whether pre-service teacher education courses are preparing students for working in datafied system is unknown, but until they do, datafication will benefit teachers with statistical and/or data expertise or those with enough time and motivation to develop skills in these areas.

Erosion of Student’s Digital Privacy

The final category is the *erosion of student’s digital privacy*. While it might seem an obvious point, many of the digital products in schools are owned by third-party companies, like Alphabet, Microsoft, and Adobe. These digital products are not created or owned by schools or departments and are often commercial in nature. While this is not necessarily an issue, the bottom line is they are businesses that need to make money (even if their overarching goal is to help educate children and young people). Many EdTech companies rely on subscriptions; however, selling data as a second stream of income has become common practice (Kelly et al. 2023). Personal data can be aggregated and processed, enabling detailed profiles of children and young people to be constructed.

Human Rights Watch (2022) reported many EdTech companies engaging in practices that centre around the collection and use of vast amounts of children’s data—often

beyond what is necessary to provide the product or service. In doing so, they undermine or put at risk children's rights, including their right to privacy. This data can be used to target children and young people, or their families, for commercial and other purposes. The realities of how education data is currently processed and used stand in stark contrast to what is considered best practice when it comes to children's rights. For example, Article 16 of the United Nations Convention on the Rights of the Child (UN CRC) states: 'No child shall be subjected to arbitrary or unlawful interference with his or her privacy.' (United Nations 1990) In addition, parents are opposed to their children using EdTech products that exploit their data for profiling and targeting purposes.

Conclusion

Of the four data harms discussed here, three directly impact children and young people, while the fourth adds additional burdens to teachers and school staff. Children and young people's education data should only be used in ways that benefit their learning. However, it seems that the neoliberal logics that underpin contemporary education and the problematic business model of many EdTech companies are making this increasingly difficult. This will only be exacerbated with the continued embrace of emerging and unregulated technologies such as AI in Australia (Cassidy 2024), the US (Office of Education Technology 2023), and the UK (Acres 2023).

It is high time that EdTech companies, including BigTech, address the concerns of parents and families and are brought into line with what the United Nations and other international and national organisations are recommending. With 196 countries having signed the UN CRC (United Nations 1990), this should be a baseline expectation of EdTech internationally. This may run counter to the neoliberal logics that underpin the increasing and expanding datafication of schooling; however, it will ensure that the harms detailed above are minimised.

Little attention has been paid to the affective dimensions of data tracking and monitoring that takes place through EdTech platforms; however, some students are already reporting a sense of powerlessness in relation to EdTech platforms. It is therefore important that we ask what impact this might have on their sense of agency and trust in the long term. While data-driven systems may have become an inevitable feature of education, it is essential that we reveal the negative implications to those that work in schools and education authorities so that we can advocate for change.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acres, T. (2023). AI could be used to provide 'personalised learning' to schoolchildren, Rishi Sunak says. Sky News, 12 June. <https://news.sky.com/story/ai-could-be-used-to-provide-personalised-learning-to-schoolchildren-rishi-sunak-says-12900840>. Accessed 13 February 2024.
- Ball, S. J. (2003). The Teacher's Soul and the Terrors of Performativity. *Journal of Education Policy*, 18(2), 215–228. <https://doi.org/10.1080/0268093022000043065>
- Cassidy, C. (2024). ChatGPT is Coming to Australian Schools. Here's What You Need to Know. The Guardian Australia, 23 January. <https://www.theguardian.com/australia-news/2024/jan/23/chatgpt-in-australian-schools-what-you-need-to-know-law-changes>. Accessed 3 February 2024.
- Clutterbuck, J., Hardy, I., & Creagh, S. (2021). Data Infrastructures as Sites of Preclusion and Omission: The Representation of Students and Schooling. *Journal of Education Policy*, 38(1), 93–114. <https://doi.org/10.1080/02680939.2021.1972166>.
- Connell, R. (2013). The Neoliberal Cascade and Education: An Essay on the Market Agenda and its Consequences. *Critical Studies in Education*, 54(2), 99–112. <https://doi.org/10.1080/17508487.2013.776990>.
- Couldry, N., & Mejias, U. A. (2019). *The Costs of Connection: How Data is Colonizing Human Life and Appropriating it for Capital*. Redwood City, CA: Stanford University Press.
- Dixon-Román, E., Nichols, T. P., & Nyame-Mensah, A. (2020). The Racializing Forces of/in AI Educational Technologies. *Learning, Media and Technology*, 45(3), 236–250. <https://doi.org/10.1080/17439884.2020.1667825>.
- Human Rights Watch. (2022). 'How Dare They Peep Into My Private Life?' Children's Rights Violations by Governments That Endorsed Online Learning During the Covid-19 pandemic. <https://www.hrw.org/report/2022/05/25/how-dare-they-peep-my-private-life/childrens-rights-violations-governments>. Accessed 20 January 2024.
- Jarke, J., & Breiter, A. (2019). Editorial: the datafication of education. *Learning, Media and Technology*, 44(1), 1–6. <https://doi.org/10.1080/17439884.2019.1573833>.
- Kelly, G., Graham, J., & Garton, S. (2023). State of Kids' Privacy: Who is Monetizing Our Data? A General Lack of Transparency Leads to a Confusing Landscape. Common Sense Media Report. <https://www.common SenseMedia.org/research/2023-state-of-kids-privacy>. Accessed 12 February 2024.
- Lu, A., Ackerman, M., Marcu, G., & Dillahunt, T. (2021). Coding Bias in the Use of Behaviour Management Technologies: Uncovering Sociotechnical Consequences of Data-Driven Surveillance in Classrooms. In *DIS '21: Proceedings of the 2021 ACM Designing Interactive Systems Conference*. New York: Association for Computing Machinery. <https://doi.org/10.1145/3461778.3462084>.
- Manolev, J., Sullivan, A., & Slee, R. (2019). The Datafication of Discipline: ClassDojo, Surveillance and a Performative Classroom Culture. *Learning, Media and Technology*, 44(1), 36–51. <https://doi.org/10.1080/17439884.2018.1558237>.
- Noble, S. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.
- Office of Education Technology. (2023). Future of Teaching and Learning: Insights and Recommendations. Washington, DC: Office of Education Technology. <https://tech.ed.gov/ai-future-of-teaching-and-learning/>. Accessed 14 of February 2024.
- Pangrazio, L., & Sefton-Green, J. (2024). Digital Literacies as a 'Soft power' of Educational Governance. In B. Williamson, J. Komljenovic, & K. N. Gulson (Eds.), *World Yearbook of Education 2024: Digitalisation of Education in the Era of Algorithms, Automation and Artificial Intelligence*. Abingdon, Oxon, UK: Routledge.
- Pangrazio, L., Selwyn, N., & Cumbo, B. (2023a). A Patchwork of Platforms: Mapping the Digital Infrastructures of Schools. *Learning, Media and Technology*, 48(1), 65–80. <https://doi.org/10.1080/17439884.2022.2035395>.
- Pangrazio, L., Selwyn, N., & Cumbo, B. (2023b). Tracking technology: exploring student experiences of school datafication. *Cambridge Journal of Education*, 53(6), 847–862. <https://doi.org/10.1080/0305764X.2023.2215194>.
- Perrotta, C. (2023). Advancing Data Justice in Education: Some Suggestions Towards a Deontological Framework. *Learning, Media and Technology*, 48(2), 187–199. <https://doi.org/10.1080/17439884.2022.2156536>.

- Perrotta, C., Gulson, K. N., Williamson, B., & Witzemberger, K. (2021). Automation, APIs and the Distributed Labour of Platform Pedagogies in Google Classroom. *Critical Studies in Education*, 62(1), 97–113. <https://doi.org/10.1080/17508487.2020.1855597>.
- Selwyn, N. (2021). The Human Labour of School Data: Exploring the Production of Digital Data in Schools. *Oxford Review of Education*, 47(3), 353–368. <https://doi.org/10.1080/03054985.2020.1835628>.
- Selwyn, N., Nemorin, S., & Bulfin, S. (2018). *Everyday Schooling in the Digital Age*. London and New York: Routledge.
- Selwyn, N., Pangrazio, L., & Cumbo, B. (2022). Data Classes: An Investigation of the People That ‘Do Data’ in Schools. In L. Pangrazio & J. Sefton-Green (Eds.), *Learning to Live with Datafication: Educational Case Studies and Initiatives from Around the World*. Abingdon Oxon, UK: Routledge.
- United Nations. (1990). Convention on the Rights of the Child. <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-child>. Accessed 6 March 2024.
- Williamson, B. (2019). Datafication of education: A critical approach to emerging analytics technologies and practices. In H. Beethem & R. Sharpe (Eds.), *Rethinking Pedagogy for a Digital Age Principles and Practices of Design* (pp. 212–226). London and New York: Routledge.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Public Affairs.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.