Centering Student Voice in Developing Learning Analytics and Artificial Intelligence–Embedded Supports

Emily Oakes, Yih Tsao, and Victor Borden

About the Authors

Emily Oakes is the Principal Unizin IT Consultant and University Data Steward for Learning Management and Learning Analytics Data for Indiana University. Yih Tsao is a project research associate with Indiana University's Charting the Future initiative. Victor Borden is professor of higher education and project director in the Indiana University Center for Postsecondary Research.

Abstract

Accelerating advancements in learning analytics and artificial intelligence (AI) offers unprecedented opportunities for improving educational experiences. Without including students' perspectives, however, there is a potential for these advancements to inadvertently marginalize or harm the very individuals these technologies aim to support. This article underscores the risks associated with sidelining student voices in decision-making processes related to their data usage. By grounding data use within a social justice framework, we advocate for a more equitable and holistic approach. Drawing on previous research as well as insights we have gathered from a student panel, we outline effective methods to integrate student voices. We conclude by emphasizing the long-term implications for the institutional research field, arguing for a shift toward more inclusive and student-centric practices in the realm of learning analytics and AI-embedded supports.

The AIR Professional File, Fall 2023

Article 162

https://doi.org/10.34315/apf1622023

INTRODUCTION: ADVANCEMENTS IN ANALYTICS AND ARTIFICIAL INTELLIGENCE

Institutional research (IR) professionals have become increasingly central to college and university efforts to improve student success through the use of empirical research and reporting. This tradition goes back to the early 20th century when information technologies and statistical methods were relatively cumbersome, through the information technology explosion of the late 20th century when tools like personal computers, spreadsheets, and statistical packages allowed for more-rapid deployment of research results. The 20-plus years since the beginning of the new millennium have seen another explosion of capacity, with institutional data supplemented by diffuse information systems available from national data systems that can be used for benchmarking and tracking students from their early school years, through college, and into the workforce. Officials at many colleges and universities have had great success leveraging such data systems, as countless sessions at the annual Association for Institutional Research (AIR) Forum have demonstrated.

Recent advances in predictive analytics have opened new possibilities in providing direct support to students—to the instructors who teach them, to the advisors who support them, and to many other new types of professionals that have roles in helping students navigate the increasingly complicated choices available to them within a particular institution and across the higher education landscape. Artificial intelligence (AI) now offers a quantum leap in capabilities that students, faculty, and staff can leverage to support student learning and success. However, there is much peril along with the promise of these technologies: instructors cannot easily tell whether the work submitted by students represents solely their own thinking or if it was aided by AI. It has been demonstrated, too, that AI can contribute to widening equity gaps due to bias inherent in algorithms as well as to equity gaps in access to and use of this powerful technology (Ahn, 2022; Alonso et al., 2020).

While some tremendous successes have already been realized, there are incalculable opportunities still to be discovered. Critical to the discovery of those opportunities is ensuring the involvement of the voice of our most important population: students. An oft-cited achievement in the use of institutional data is Georgia State University's (GSU) predictive analytics service. Since partnering with EAB in 2012, GSU has seen its graduation rates increase by more than 35 percentage points; as of 2023 those rates have been consistent across racial and ethnic lines for 7 years. The institution has increased degrees awarded by 84% and more than doubled the number awarded to low-income and minority students. Powering their alerts are 10 years of data that were reviewed to identify 800 factors that correlate with challenges completing their degrees on time (Calhoun-Brown, 2023). Of equal importance, 42 advisors were hired alongside the service's launch, enabling more advisor–student interactions (Kurzweil & Wu, 2015). GSU has profoundly and positively impacted its students' paths to success, as have many other institutions, aided by the use of advanced information and analytic capacities.

But, as noted, GSU's successes involved more than just leveraging new analytic technologies. The institution was already seeing consistent improvements in its graduation rates before the implementation of its advising alert system in 2012

(GSU, 2021). In their 2015 case study, Kurzweil & Wu (2015) noted that GSU's incredible results are not related to a single solution, but rather to the institution's overall approach to problem solving. Staff members at GSU use the institution's data warehouse to find barriers to graduation and resolve those barriers through a cycle of implementing interventions to remove identified barriers; they assess their effectiveness and scale them up if they find them to be effective. As Kurzweil & Wu (2015, p. 18) note, "It is the process, and not merely its outputs, that other institutions should seek to replicate."

GSU's process included opportunities for centering the student voice. In this article we first describe considerations and risks when student voices are not included in deciding how their data will be used. Next, we discuss ways to ground data use in a social justice framework. Finally, we share perspectives and recommendations on how to support students' successes.

Although applications of AI often operate on a more diverse range of data types and use techniques that are different from predictive analytics, the issues considered in this article apply equally, if not more strongly, given that that the user of AI's output is even farther removed from the analysis process than is the user of predictive analytics.

Considerations and Risks When Student Voices Are Not Included

Understanding that there are risks when students, especially students from marginalized populations, are not involved in uses of their data is critical to avoiding those risks. Fortunately, many lessons have already been learned regarding a lack of participation in data use generally that institutional researchers can consider in the context of their work as they move forward in deploying AI as part of their information use strategies.

First, concerns have been raised among scholars and practitioners working toward data justice that data reflect social ideas of the default as implicitly defined by those with power in a particular context: White, heterosexual, cisgender, abled, neurotypical, financially comfortable, and so on (Benjamin, 2019; D'Ignazio & Klein, 2020). When data are captured, structured, interpreted, and applied on the assumption of a particular default, those who fall outside of that category are less likely to benefit and more likely to possibly experience harm.

Consider AI researcher, artist, and advocate Joy Buolamwini's now-famous experience discovering bias in facial analysis software (Kantayya, 2020). While interacting with the software, Buolamwini found that the software was unable to identify her darker-skinned face (a label that itself implies a default), despite successfully capturing her lighter-skinned colleagues' faces. The software was similarly able to identify the features of a plain white mask she placed over her own face (Kantayya, 2020). Buolamwini and computer scientist Timnit Gebru had previously found that multiple data sets used to train facial recognition software had included majority lighter-skinned subjects, causing the software to frequently misclassify darker-skinned faces, with the greatest number of errors occurring when the software attempted to analyze the faces of darker-skinned women (Buolamwini & Gebru, 2018). These issues of bias and unfairness occur with generative AI, such as Chat Generative Pre-trained Transformer (ChatGPT), as well, and thus require training users on diverse data, careful monitoring, and other bias mitigation tactics (Kasneci et al., 2023). Mitigation strategies should be defined in use policies informed by impacted populations (i.e., students of various identities) to surface issues that others outside those populations may not be aware of. As institutions invest in data-powered identity-based outreach, AI video assessment and proctoring, AI-assisted admissions, or staff interview software, and so on, their similar investments in mitigation strategies will only grow in importance.

Early alert systems are a useful tool for demonstrating the practical risks when services do not incorporate student-guided use policy. Early alert systems are frequently implemented in higher education in an effort to increase retention (Parnell et al., 2018). These systems use data about students that are based on some predefined metrics to identify when students are at greater risk of incurring negative academic consequences, and send an alert to instructors or academic support staff so that they may intervene as appropriate (Hanover Research, 2014). Interventions might include offering tutoring, having a student meet with an advisor, assigning a mentor, or referring a student to a relevant social service (Ekowo & Palmer, 2017).

Numerous risks arise when a diversity of student voices have not been considered in the development, deployment, and operation of early alert systems. First, the integration of multiple data sets means that a risk label can be made more broadly visible, which creates opportunities for riskiness to be assumed in contexts unrelated to the one that the risk was measured against in the first place (Benjamin, 2019; Prinsloo & Slade, 2016). This is additionally problematic given that student identities and circumstances frequently change: while data about students often tend to be rigid, the realities of their lives are not (Slade & Prinsloo, 2013). Without an opportunity to dispute or otherwise provide narrative context alongside their data, circumstances perceived as negative and permanently recorded by an institution official can follow students throughout their academic careers.

In their review of relevant literature, Braunack-Mayer et al. (2020) found that students have expressed concern across multiple studies about being labeled “at risk”; these authors note that being categorized in certain ways could bias their instructors such that they exclude the categorized students from future academic opportunities. In this way, the label “risky” becomes a quality inherent to a student, detached from its use as a descriptor applied to those who are being failed by a specific process or system. Nopper (2019, p. 170) refers to the “digital profile assessed to make inferences regarding character in terms of credibility, reliability, industriousness, responsibility, morality, and relationship choices” as “digital character” that is used to paternalistically “help” individuals, often without their knowledge or consent. (See also Braunack-Mayer et al., 2020.) This focus on applying interventions based on a student’s digital character situates them as data objects or passive recipients of services rather than as autonomous agents (Kruse & Pongsajapan, 2012; Prinsloo & Slade, 2016; Roberts et al., 2016; Rubel & Jones, 2014). Given that groups of students have also expressed such concerns about threats to their autonomy by these systems themselves, it is critical that they are provided mechanisms for having their voices considered (Roberts et al., 2016). This example is not intended to imply that all early alert systems are problematic—there is evidence that students do consider them beneficial (Atif et al., 2015; Roberts et al., 2016). Rather, the example is used here to illustrate the potential issues that may arise if development of such systems is not aligned with student-informed policies for use.

Grounding Data Use in a Social Justice Framework

To productively address risks like those described, we suggest that higher education officials align their efforts to grow data capacities and use AI-infused solutions with their diversity, equity, and inclusion priorities. This is not a novel approach to data use: the social impacts of mass data use have received increasing attention for more than a decade. In 2012, Facebook gained significant media attention around its nonconsensual research on and manipulation of users' moods; the use of its data by political consulting firm Cambridge Analytica in 2018 helped raise public consciousness about mass data's capabilities and misuses (Meyer, 2014; Zialcita, 2019). Zuboff (2019) described how surveillance capitalism—the widespread collection and commodification of personal data by corporations—poses significant threats to society, privacy, and autonomy. Relatedly, O'Neil (2016) laid out numerous examples of the harm Big Data algorithms can cause across contexts, including their use in college rankings and teacher evaluations, and Wachter-Boettcher (2017) discussed the lack of diversity and inclusivity in the technology industry, leading to sexist, inaccessible, and otherwise biased systems. Additionally, Noble (2018) detailed the ways that search engines reinforce racism, sexism, and other forms of oppression; Benjamin (2019) broadened Noble's work, discussing additional applications of data that cause harm to vulnerable populations, including in AI systems.

Applications of data and the calculations we apply to data (i.e., algorithms) have been investigated from a variety of perspectives and within numerous contexts. Out of these investigations has developed the concept of data justice—a framework for engaging with the ways datafication and society intersect with an explicit social justice focus. While there are diverse approaches to and definitions of data justice, there are some themes, including the recommendation to meaningfully collaborate with the individuals whose data will be captured and used during the conception, development, and implementation of data-based systems and the policies governing them (Dencik et al., 2019; Dencik & Sanchez-Monedero, 2022). In academia, these individuals are often our students.

In the remainder of this article, we consider the implications of using a social justice framework for advancing the use of generative AI and other Big Data applications within higher education institutions. This framework derives from a focus on minoritized populations, such as Indigenous peoples and other racial/ethnic minorities, who are often underrepresented within postsecondary institutions. We believe, however, that the ideas pertain more generally to students who, although often the largest group of constituents of a college or university, are not consulted about the use of their personal data within such applications.

PERSPECTIVES AND RECOMMENDATIONS

Numerous communities have shared their perspectives on and recommendations for data use as it relates to their unique experiences. While these communities are not monolithic, the concerns they raise reflect themes that might otherwise go unidentified by those who develop and deploy AI and Big Data applications (D'Ignazio & Klein, 2020).

One such group advocating for data justice is the Native Nations Institute (NNI). The NNI defines a Native nation's data as "any facts, knowledge, or information about the nation and about its citizens, lands, resources, programs, and communities. Information ranging from demographic profiles to educational attainment rates, maps of sacred lands, songs, and social media activities are all data" (Rainie et al., 2017, p. 1). The NNI aims to promote Indigenous data sovereignty using the CARE Principles for Indigenous Data Governance that were developed by the Research Data Alliance's International Indigenous Data Sovereignty Interest Group in 2018 and published in 2020 (Carroll et al., 2020). The CARE Principles and their subcomponents are summarized in Table 1.

Table 1. The CARE Principles for Indigenous Data Governance

Principle	Component
Collective Benefit	For inclusive development and innovation For improved governance and citizen engagemen For equitable outcomes
Authority to Control	Recognizing rights and interests Data for governance Governance of data
Responsibility	For positive relationships For expanding capability and capacity For Indigenous languages and worldviews
Ethics	For minimizing harm and maximizing benefits For justice For future use

Source: Adapted from Carroll et al., 2020, Figure 2.

The Responsibility principle's first subsection, "For positive relationships," identifies that "Indigenous data use is unviable unless linked to relationships built on respect, reciprocity, trust, and mutual understanding, as defined by the Indigenous Peoples to whom those data relate" (Carroll et al., 2022, p. 4). The following subsections, "For Expanding Capability and Capacity" and "For Indigenous Languages and Worldviews," require efforts to increase data literacy and to ground data in the world views and the lived experiences of Indigenous peoples, respectively. Each of these subsections implies some form of collaboration between institution officials using Indigenous students' data and the students themselves: to create mutual understanding, to increase data literacy between both parties, and to enable Indigenous students to (consensually) share their experiences.

When considering the use of early alert systems, it is important to note that the CARE Principles for Indigenous Data Governance specify that ethical data not portray Indigenous peoples in terms of deficit, and that benefits and harms should be evaluated from the perspective of the Indigenous peoples the data relate to (Carroll et al., 2020). This guidance provides a model for data use policy development that may be applied to other student populations regardless of identity; rather than administrators determining what may harm or benefit communities, administrators can consult with those communities to provide their contextualized view of potential risks and benefits, and to describe assets to highlight with students.

Although designed with specific focus on a highly marginalized population, the principles can be applied more generally to incorporating student voice into the formulation of machine learning (ML), AI, and other Big Data applications and resources. However, these principles also remind us that we need to pay special attention to the voices of marginalized student populations, such as racially minoritized students and other subgroups that are not well represented by the dominant student culture.

Other issues related to data capture have been identified as well. Ruberg & Ruelos (2020) note that it is difficult to accurately represent gender and sexuality using traditional demographic capture-and-reporting techniques. Those authors provide multiple recommendations based on their findings: (1) When capturing gender and sexuality, multiple answer possibilities should be available. (2) Gender and sexuality identities may change, and all reported identities are valid unless the individual states otherwise. (3) Collaboration with relevant communities is critical for understanding and accurately capturing their identities.

Finally, marginalized groups are often centered and surveilled by both punitive and purportedly supportive systems, which promotes feelings of threatening hypervisibility (Benjamin, 2019). Asher et al.'s (2022) survey of student perspectives on library analytics found that students in minority racial/ethnic groups and those of lower socioeconomic status were more concerned than the overall student population about the privacy of their personal data, thus supporting this perspective in the academic context (Asher et al., 2022). Collaborating with students, especially those who experience heightened surveillance, may help to shift support methods such that students experience them in a less threatening manner. To this point, GSU's predictive advising service provides another example: risk factors are shared with students as well as with advisors, promoting transparent conversations; and advisors are thoroughly trained on how to use the service as well as how to have discussions about its outputs with students (Bailey et al., 2019).

Methods for Including Student Voices

There are a variety of potential methods for involving student perspectives when developing access and use policies. West et al. (2020) note that these methods could include research into students' descriptions of their own needs, concerns, and ideas for how learning analytics might benefit them, as well as the creation of user users' stories and principles against which data-based tools may be built. Jones et al. (2019, 2020, 2023) demonstrate adaptable methods for gathering student feedback in their studies by collecting student perspectives in three phases across 3 years: first, they conduct interviews with undergraduate students across eight U.S. institutions, then they send a quantitative survey to random samples of students across the same eight institutions, and finally they hold virtual focus groups centered on discussions of data use scenarios rooted in real-life practice. Data for Black Lives' report, *Data Capitalism + Algorithmic Racism* (Milner & Traub, 2021), suggests a few methods for supporting collective data practice that can be adjusted for the higher education context, such as Collington's (2019) proposed "system including a digital platform for debating and deciding priorities for use of public data" (Milner & Traub, 2021, p. 26).

An even more-robust strategy is provided in *A Toolkit for Centering Racial Equity Throughout Data Integration from Actionable Intelligence for Social Policy*, which includes guidance for involving community voices at every stage of design, use, and implementation of data-infused practices (Hawn Nelson et al., 2020). While the *Toolkit* was developed to support those using data for civic purposes, many of its recommendations apply to higher education data uses and align with calls from the learning analytics field to include student voices at all levels of data use, from design development through membership in oversight committees (Braunack-Mayer et al., 2020). Among other recommendations, the *Toolkit* suggests involving diverse community members in discussions about algorithms and their purposes early in the design stage, inviting people with multiple perspectives to provide potential interpretations of data that will be used.

Using a Student Panel Methodology to Center Student Voice

A method that incorporated both surveys and focus groups was devised as part of a university-wide student success initiative within the authors' institution. Fifteen students were initially recruited from across the institution's seven campuses, and most of the same students attended each panel, which helped to establish an environment of open sharing. For the panel exploring student views on the use of learning analytics and Big Data, the student panelists first reviewed a set of materials related to the use of learning analytics at several different universities, as well as among a community of learning management system users. Students then completed a survey including questions about their awareness of the types of data collected, about their privacy and agency regarding learning data, about issues related to instructors and advisors who have access to and use the data, as well as questions about the benefits and risks with the use of these data. Student responses were split somewhat evenly on the awareness of the types of data that were being collected, but the majority (70%) of the students disagreed with the statement that they were adequately informed about how their data were being used. Interestingly, while more than 80% of the students agreed that there are benefits to making these data available to their instructors, 40% agreed with the reverse statement that such awareness may also negatively impact their motivation and engagement in a course.

The panel discussion focused on four questions for which the students used Google's Jamboard app to record and organize their ideas into themes. The four questions asked were the following:

1. What were your reactions to learning about the kind of learning data that your instructors can access?
2. What were your reactions to advisors' use of Early Alert Systems?
3. How do you feel about your learning data being used to identify that you are struggling in a course?
4. What would you like your instructor to communicate to you about learning data use in your courses?

After completing the analysis, the students were split into two groups to formulate a plan or list of recommendations regarding safeguards/procedures that should be in place to ensure that inequities or biases are not introduced in the use of learning data in a course. Table 2 shows an organization of the thematic responses to this task from the student panelists.

Table 2. Thematic Responses from the Breakout Room Activity During the Panel

Themes	Examples/Explanations
Possible forms of biases in current practices	• Instructor shows favoritism for students struggling less. • Not all struggling students receive the appropriate outreach. • There are biases regarding students' socioeconomic status backgrounds. • Student backgrounds (e.g., they were homeschooled, are first-generation students) lead to different knowledge or resources used to reach out to students with invisible needs.
Theme 1 **Transparency/Open Communication**	• Student consent should be collected before the data are collected and shared with instructors, advisors, or any other parties. • The types of data collected or shared should be communicated clearly to both students and instructors. • Researchers should explain to the students how the data are being used or will be used. • Students should have access to their own learning data. • All students should have equal access to resources and support.
Theme 2 **Training**	• Instructors, advisors, and anyone who may be in close contact with any student data should receive bias and diversity training. • Instructors and advisors should be trained in how to be sensitive to when and especially how to reach out to struggling students with more care and attention to their words. • Instructors should be trained in how to initiate contact with students.
Theme 3 **Human Oversight**	• There should be a separate office that analyzes student data before the data are used by instructors or advisors for reaching out to students, or by students themselves. • There should be more communications or surveying of students to better understand their perspectives and opinions. • Teachers and administrators or advisors should be allowed to review their decisions based on their bias trainings.

Through the survey responses and the panel discussions, student data use is clearly a topic that is sensitive and requires more attention to its ethics and to the treatment of individuals. When using AI and Big Data in higher education, we must be more diligent in protecting the humans behind the numbers. Students may feel uncomfortable when they become aware of the data that are being collected about them; that sense of discomfort can escalate when the data are shared outside of the context where they are generated, such as in-class data being shared with an academic advisor. Finally, the panel discussion revealed a concern about how students are treated when the data are used: Will they be treated fairly? Is outreach done with sensitivity and care? And how can marginalization and biases be avoided in terms of access to resources and support?

This student panel methodology serves to center student voice in IR and to inform policies. To accurately represent students' voices, however, it is essential to reflect the diversity of the student body to avoid bias. For example, while this student panel was recruited from various campuses of the same institution, more than half of the student panelists were from the main campus. Even though this accurately reflects the representation of students across the university, it skews the possible viewpoints and practices experienced by the students. Similarly, their classification (year), gender, race/ethnicity, socioeconomic status, and other demographics should also be taken into consideration when recruiting to prevent representation disparity in data that could lead to unjust applications, such as Buolamwini's facial analysis software, as mentioned before (Buolamwini & Gebru, 2018).

THE IMPORTANCE OF BRINGING IN STUDENT VOICES AND IMPLICATIONS FOR THE INSTITUTIONAL RESEARCH FIELD

Actionable Intelligence for Social Policy's *Toolkit* (Hawn Nelson et al., 2020), discussed earlier, recommends questioning how data use can help communities (i.e., students, in our context) to interrogate systems, as opposed to using data only to identify how to treat those communities. To align with effective and ethical practice, we recommend that institutional researchers intentionally and continually frame their work as student-centric as opposed to intervention-centric, and that they direct their actions in response to collaborations with students primarily toward the systems the students interact with instead of the students themselves (Hawn Nelson et al., 2020; Kruse & Pongsajapan, 2012; Slade & Prinsloo, 2013). Actionable Intelligence for Social Policy's *Toolkit* (Hawn Nelson et al., 2020) includes activities that may be adapted for this purpose. Practical steps for operationalizing racial equity in data use are included, as well as numerous real-world examples of the guidance in practice. Once again, GSU's approach to data use in support of student success provides an example of this practice in action: by asking first whether the institution is the problem (i.e., by interrogating the institution's systems), GSU has been able to find and resolve significant barriers facing students (Kurzweil & Wu, 2015; Zipper, 2022). It is crucial to involve student voices: in addition to data, students can provide context for *why* something was a barrier as well as advice for how institutions can break down barriers.

It is critical that student voices are actively centered when developing data access and use policies. When we authentically include student voices in the development of data policy, we can uncover novel opportunities that will be situated in the contexts of our most important constituents. We can learn what they value and what their challenges are from their own perspectives instead of mediated, decontextualized data sets. Including students in the development of data policy and system development increases trust, and fosters development of systems and initiatives that support success as students define it. In this article, we have shared one approach used for our context and numerous other approaches that could be adapted, and we invite institutional researchers to consider how they may take advantage of these methods for their contexts as well.

REFERENCES

Ahn, J. (2022). Exploring the negative and gap-widening effects of EdTech on young children's learning achievement: Evidence from a longitudinal dataset of children in American K–3 classrooms. *International Journal of Environmental Research and Public Health, 19*(9), 5430 (19pp.). https://www.mdpi.com/1660-4601/19/9/5430

Alonso, C., Berg, A., Kothari, S., Papageorgiou, C., & Rehman, S. (2022). Will the AI revolution cause a great divergence? *Journal of Monetary Economics, 127*, 18–37. https://www.sciencedirect.com/science/article/abs/pii/S0304393222000162

Asher, A. D., Briney, K. A., Jones, K. M., Regalado, M., Perry, M. R., Goben, A., Smale, M. A., & Salo, D. (2022). Questions of trust: A survey of student expectations and perspectives on library learning analytics. *Library Quarterly, 92*(2), 151–171. https://www.journals.uchicago.edu/doi/abs/10.1086/718605

Atif, A., Richards, D., & Bilgin, A. (2015). Student preferences and attitudes to the use of early alerts. In T. Bandyopadhyay, D. Beyene, & S. Negash (Eds.), *2015 Americas Conference on Information Systems* (pp. 1–14). Americas Conference on Information Systems. https://researchers.mq.edu.au/en/publications/student-preferences-and-attitudes-to-the-use-of-early-alerts

Bailey, A., Vaduganathan, N., Henry, T., Leverdiere, R., & Jacobson, M. (2019). *Turning more tassels: How colleges and universities are improving student and institutional performance with better advising*. Boston Consulting Group and NASPA Student Affairs Administrators in Higher Education. https://web-assets.bcg.com/img-src/BCG-Turning-More-Tassels-Jan-2019_tcm9-212400.pdf

Benjamin, R. (2019). Assessing risk, automating racism: A health care algorithm reflects underlying racial bias in society. *Science, 366*(6464), 421–422. https://www.science.org/doi/10.1126/science.aaz3873

Braunack-Mayer, A. J., Street, J. M., Tooher, R., Feng, X., & Scharling-Gamba, K. (2020). Student and staff perspectives on the use of Big Data in the tertiary education sector: A scoping review and reflection on the ethical issues. *Review of Educational Research, 90*(6), 788–823. https://journals.sagepub.com/doi/abs/10.3102/0034654320960213?journalCode=rera

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In S. A. Friedler and C. Wilson (Eds.), *Proceedings of the 1st Conference on Fairness, Accountability and Transparency (Proceedings of Machine Learning Research), Journal of Machine Learning Research, 81*, 77–91. http://proceedings.mlr.press/v81/buolamwini18a.html

Calhoun-Brown, A. (2023, January 9). How data and technology can improve advising and equity. *Chronicle for Higher Education*. https://www.chronicle.com/article/how-data-and-technology-can-improve-advising-and-equity

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE principles for Indigenous data governance. *Data Science Journal*, 19(1), 43. https://doi.org/10.5334/dsj-2020-043

Carroll, S. R., Garba, I., Plevel, R., Small-Rodriguez, D., Hiratsuka, V. Y., Hudson, M., & Garrison, N. A. (2022). Using Indigenous standards to implement the CARE principles: Setting expectations through tribal research codes. *Frontiers in Genetics, 13*, 823309. https://www.frontiersin.org/articles/10.3389/fgene.2022.823309/full

Collington, R. (2019). Digital public assets: Rethinking value and ownership of public sector data in the platform age. Common Wealth. https://www.common-wealth.org/publications/digital-public-assets-rethinking-value-access-and-control-of-public-sector-data-in-the-platform-age

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press. https://mitpress.mit.edu/9780262547185/data-feminism/

Dencik, L. & Sanchez-Monedero, J. (2022). Data justice. *Internet Policy Review, 11*(1). https://policyreview.info/articles/analysis/data-justice

Dencik, L., Hintz, A., Redden, J., & Treré, E. (2019). *Exploring data justice: Conceptions, applications and directions. Information, Communication & Society, 22*(7), 873–881. https://www.tandfonline.com/doi/pdf/10.1080/1369118X.2019.1606268

Ekowo, M., & Palmer, I. (2017). *Predictive analytics in higher education*. New America. https://kresge.org/wp-content/uploads/2020/06/predictive-analytics-guidingprinciples.pdf

Georgia State University (GSU). (2021). Strategic plan goal 1: Student success. Retrieved June 21, 2023, from https://strategic.gsu.edu/accomplishments/goal-1-accomplishments/

Hanover Research. (2014, November). *Early alert systems in higher education*. https://www.hanoverresearch.com/wp-content/uploads/2017/08/Early-Alert-Systems-in-Higher-Education.pdf

Hawn Nelson, A., Jenkins, D., Zanti, S., Katz, M., Berkowitz, E., Burnett, T., & Culhane, D. (2020). A toolkit for centering racial equity throughout data integration. Actionable Intelligence for Social Policy. https://aisp.upenn.edu/resource-article/a-toolkit-for-centering-racial-equity-throughout-data-integration/

Jones, K. M. L., Asher, A., Goben, A., Perry, M. R., Salo, D., Briney, K. A., & Robertshaw, M. B. (2020). "We're being tracked at all times": Student perspectives of their privacy in relation to learning analytics in higher education. *Journal of the Association for Information Science and Technology, 71*(9), 1044–1059. https://doi.org/10.1002/asi.24358

Jones, K. M. L., Goben, A., Perry, M. R., Regalado, M., Salo, D., Asher, A. D., Smale, M. A., & Briney, K. A. (2023). Transparency and consent: Student perspectives on educational data analytics scenarios. *Portal: Libraries & the Academy, 23*(3), 485–515. https://muse.jhu.edu/article/901565

Jones, K. M. L., Perry, M. R., Goben, A., Asher, A., Briney, K. A., Robertshaw, M. B., & Salo, D. (2019). In their own words: Student perspectives on privacy and library participation in learning analytics initiatives. [Paper presentation]. ACRL 19th National Conference, "Recasting the Narrative," Cleveland, OH. https://alair.ala.org/handle/11213/17656

Kantayya, S. (Director). (2020). *Coded bias* [Documentary]. 7th Empire Media. https://www.youtube.com/watch?v=jZl55PsfZJQ (trailer).

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasserm, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*(April). https://www.sciencedirect.com/science/article/pii/S1041608023000195?via%3Dihub

Kruse, A., & Pongsajapan, R. (2012). Student-centered learning analytics. *CNDLS Thought Papers, 1*(9), 98–112. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=692c8661e4c13efe3559cce96694689ce72f307e

Kurzweil, M., & Wu, D. (2015, April 23). Building a pathway to student success at Georgia State University. Ithaka S+R. https://sr.ithaka.org/publications/building-a-pathway-to-student-success-at-georgia-state-university/

Meyer, M. N. (2014, June 30). Everything you need to know about Facebook's controversial emotion experiment. *Wired*. https://www.wired.com/2014/06/everything-you-need-to-know-about-facebooks-manipulative-experiment/

Milner, Y., & Traub, A. (2021). *Data capitalism + algorithmic racism*. Data for Black Lives. https://datacapitalism.d4bl.org/documents/Demos_Data_Capitalism_Final.pdf

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press. https://psycnet.apa.org/record/2018-08016-000

Nopper, T. K. (2019). Digital character in "The scored society." In R. Benjamin (Ed.), *Captivating technology: Race, carceral technoscience, and liberatory imagination in everyday life* (pp. 170–187). Duke University Press. https://www.degruyter.com/document/doi/10.1515/9781478004493-010/html?lang=en

O'Neil, C. (2016). *Weapons of math destruction: How Big Data increases inequality and threatens democracy*. Crown. https://archive.org/details/weapons-of-math-destruction_202209

Parnell, A., Jones, D., Wesaw, A., & Brooks, D. C. (2018). Institutions' use of data and analytics for student success. *EDUCAUSE: Center for Analysis and Research*. https://eric.ed.gov/?id=ED588396

Prinsloo, P., & Slade, S. (2016). Student vulnerability, agency, and learning analytics: An exploration. *Journal of Learning Analytics, 3*(1), 159–182. https://learning-analytics.info/index.php/JLA/article/view/4447

Rainie, S. C., Rodriguez-Lonebear, D., & Martinez, A. (2017). Policy brief (ver. 2): Data governance for native nation rebuilding. Native Nations Institute. https://nnigovernance.arizona.edu/policy-brief-data-governance-native-nation-rebuilding

Roberts, L. D., Howell, J. A., Seaman, K., & Gibson, D. C. (2016). Student attitudes toward learning analytics in higher education: The Fitbit version of the learning world. *Frontiers in Psychology, 7*. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01959/full

Rubel, A., & Jones, K. M. (2016). Student privacy in learning analytics: An information ethics perspective. *The Information Society, 32*(2), 143–159. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2533704

Ruberg, B., & Ruelos, S. (2020). Data for queer lives: How LGBTQ gender and sexuality identities challenge norms of demographics. *Big Data & Society, 7*(1). https://doi.org/10.1177/2053951720933286

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*(10), 1510–1529. https://doi.org/10.1177/0002764213479366

Wachter-Boettcher, S. (2017). *Technically wrong: Sexist apps, biased algorithms, and other threats of toxic tech*. W. W. Norton & Company. https://dl.acm.org/doi/10.5555/3307044

West, D., Luzeckyj, A., Toohey, D., Vanderlelie, J., & Searle, B. (2020). Do academics and university administrators really know better? The ethics of positioning student perspectives in learning analytics. *Australasian Journal of Educational Technology, 36*(2), 60–70. https://doi.org/10.14742/ajet.4653

Zialcita, P. (2019, October 30). Facebook pays $643,000 fine for role in Cambridge Analytics scandal. *NPR News*. https://www.npr.org/2019/10/30/774749376/facebook-pays-643-000-fine-for-role-in-cambridge-analytica-scandal

Zipper, T. (Host). (2022, September). Are we the problem? with Dr. Tim Renick [Audio podcast]. (Season 2, Episode 3). *An Educated Guest*. Wiley University Services. https://universityservices.wiley.com/podcast-tim-renick-are-we-the-problem/

Zuboff, S. (2019). Surveillance capitalism and the challenge of collective action. *New Labor Forum, 28*(1), 10–29. https://journals.sagepub.com/doi/full/10.1177/1095796018819461