2023
Volume 1
Article 2
https://doi.org/10.59453//XLUD7002

Data management of AI-powered education technologies: Challenges and opportunities

Hassan Khosravi[1], Shazia Sadiq[1], and Sihem Amer-Yahia[2]

[1] The University of Queensland, Brisbane, Australia

[2] CNRS, Université Grenoble Alpes, Grenoble, France

The use of AI-powered educational technologies (AI-EdTech) offers a range of advantages to students, instructors, and educational institutions. While much has been achieved, several challenges in managing the data underpinning AI-EdTech are limiting progress in the field. This paper outlines some of these challenges and argues that data management research has the potential to provide solutions that can enable responsible and effective learner-supporting, teacher-supporting, and institution-supporting AI-EdTech. Our hope is to establish a common ground for collaboration and to foster partnerships among educational experts, AI developers and data management researchers in order to respond effectively to the rapidly evolving global educational landscape and drive the development of AI-EdTech.

Keywords: AI in education, data management, educational technologies

Correspondence: Hassan Khosravi, h.khosravi@uq.edu.au

Recommended citation: Khosravi, H., Sadiq, S., & Amer-Yahia, S. (2023). Data management of AI-powered education technologies: Challenges and opportunities. *Learning Letters*, *1*, Article 2. https://doi.org/10.59453//XLUD7002

Introduction

The rise of AI technology and the growing volume and complexity of student data has enabled AI in education (AIED) and AI-powered educational technologies (AI-EdTech) to play an instrumental role in improving the design and delivery of education. For students, AI-EdTech can provide a tailored learning experience, adapting to their specific strengths and weaknesses (Becker, 2017; Fadel et al., 2019; Miao et al., 2021; Sapci & Sapci, 2020). For instructors, AI-EdTech can act as a ubiquitous teaching assistant, helping them in orchestrating the classroom, grading assessments, and answering student inquiries (Martinez-Maldonado, 2016; Swiecki et al., 2022; Yang & Zhu, 2022). Additionally, educational institutions can utilise AI-EdTech for student admission, and to identify and support students who are at risk of dropping out or who are struggling in their courses (Rastrollo-Guerrero et al., 2020; Shabaninejad et al., 2022).

The absence of a common data infrastructure, alongside data collection, modelling and access challenges, has led to most existing AI-EdTech operating in silos, lacking access to comprehensive data that captures the entirety and diversity of learning experiences. Often, learners do not have access to, or are unaware of, the data captured about them and how it is being used to inform decisions about their learning. Instructors face limitations in their ability to gain insights into student learning, implement effective pedagogical interventions, and evaluate the impact of their approaches or use of AI-EdTech on student learning. Lastly, educational institutions face challenges in ethically integrating multi-sourced data from different tools to infer students' learning experiences and to provide students and educators with the necessary data literacy skills to understand and act on the data.

In this paper, we first analyse the underlying data needs of AIED, followed by an overview of representative AI-EdTech applications designed to support learners, teachers, and institutions. We then outline the main challenges and research opportunities for an AI-EdTech data infrastructure. In the remainder of the paper, we discuss the current challenges and aspirations in developing AI-EdTech to support learners, teachers, and institutions respectively and probe opportunities to address them by leveraging state-of-the-art research in data management.

Overview of AIED research

The data behind AI-EdTech is described in Figure 1 and represents learners and learning artefacts. In the following sections, we elaborate on each element of the overview.

Data and infrastructure

Under "Data and Infrastructure", "Student and curriculum records" capture individual learners' records such as their demographics which are usually provided by learners at registration time as well as information on learning material, such as artefacts, and assessment and outcome requirements. "Learning records" capture data on learners' achievements, such as grades and assessment outcomes. Finally, "Learning logs" record learners' engagement with artefacts, feedback to learners, and collaboration among learners.

A data infrastructure for education must be designed to meet the needs of various applications. Some applications are data generators and require large-scale storage and processing of learning logs; some require real-time access to data and must rely on efficient indexing techniques; some need to provide powerful primitives such as roll-ups and drill-downs; others need to infer data about learners. This calls for defining a data model to formalise, collect and infer data about learners in a scalable and privacy-preserving manner. "Data modelling" must capture the variety of data needed in AI-EdTech, as well as the machine learning (ML) models used to infer them. The data model must represent individual and collective factors from diverse learners: e.g., different demographics and skills (Rahman et al., 2015), intrinsic and extrinsic motivation (Hackman & Oldham, 1976; Kaufmann et al., 2011; Pilourdault et al., 2017; Pilourdault et al., 2018; Posch et al., 2019), fatigue (Hata et al., 2017), team affinity and critical mass (Cao et al., 2020; Roy et al., 2015).

We note that "data collection" would require significant human effort, especially when collecting learners' cognitive and emotional states, and the use of various tools to capture physiological and neurophysiological measurements (Darvishi et al., 2022). This can be addressed with two unconventional innovations: blending implicit data collection through observing students in situ and computing individual and collective factors (Amer-Yahia & Roy, 2016), and explicit data collection following established practices in Cognitive Psychology such as the Myers–Briggs Type Indicator questionnaire, an introspective self-report indicating psychological preferences on how people perceive the world and make decisions (Briggs, 1976). This data can be used by sampling strategies and active learning to train ML models that infer diverse and representative data. In addition, facilitating access to this data is necessary and hence "data access" primitives must be devised to enable privacy-preserving, efficient and responsible access to information regarding individuals or cohorts.

Figure 1: Overview of AIED research, focusing on the examination of underlying data, applications, challenges, and opportunities in data management.

Learner-supporting AI-EdTech

The use of learner-supporting AI is fast becoming popular in mainstream education (Becker, 2017; Fadel et al., 2019; Miao et al., 2021; Sapci & Sapci, 2020) where AI-EdTech is used to automate some aspects of educators' support to help students better regulate their learning, and receive various forms of guidance and feedback on their work. Of notable mention is the use of learner models and profiles that represent a student's competencies and knowledge gaps based on their performance and interactions with the educational system (Abdi et al., 2020; Abdi et al., 2021; Bull, 2020). Learner models and profiles enable students to discover their learning preferences, identify areas for improvement, set learning objectives, and initiate discussions on areas they require assistance (Barthakur et al., 2023). Perhaps the most well-studied class of learner-supporting AI-EdTech are adaptive learning systems (VanLehn, 2011) that leverage learner models to provide an efficient, effective and customised learning experience for students by capturing data on students' competencies and interaction with various learning activities, and dynamically adapting learning content to suit their individual abilities or preferences.

The development of learner-supporting AI-EdTech must be accompanied with guarantees to learners on "data custodianship" to provide them with control over their data and how it is used to enhance their learning. This can be achieved with the use of privacy-preserving Federated Learning wand cryptographic schemes to train ML models while keeping individual data locally. The key innovative idea is to extend the relational algebra with ML operations (Luo et al., 2020), in particular with Federated Learning. Under the orchestration of a central server, data owners can participate in reward computation without revealing their raw data. This will rely on secure multi-party computations and cryptographic schemes under the honest-but-curious threat model (Marcadet et al., 2022).

Learner-supporting AI-EdTech must also provide "fair decisions and recommendations" (Kizilcec & Lee, 2020) that are understandable to students (Khosravi et al., 2022) (e.g., "You are given practice item A because you encountered challenges in completing practice item B, which likewise evaluated your proficiency in skill C."). This approach aids learners to comprehend the reasoning behind decisions and recommendations, enabling them to regulate and monitor their learning more effectively. A key opportunity here for data management research is to utilise best practices from the field of Explainable Artificial Intelligence (XAI) (Gunning et al., 2019), complemented by research on long-term fairness in recommendations (Ge et al., 2021) and adaptive recommendations that account for evolving user needs over time (Azzalini et al., 2022) to produce interpretable and fair-by-design recommendations to enhance student learning. To validate recommendations, we need to design an innovative evaluation methodology that automates experimental protocols to streamline large-scale studies. Data triangulation has the potential to reconcile large-scale results from online labour platforms and smaller-scale results from educational platforms. Participants must be provided with a free and voluntary consent form that complies with relevant regulations such as General Data Protection Regulation (GDPR).

While adaptive learning systems mostly focus on individual learning, there has also been a focus on AI-EdTech to encourage "collaborative learning" and to enable learners to connect and collaborate with their peers who share similar aspirations and interests. The common aspects of collaborative learning supported by AI can be categorised into the following: (1) group formation or peer recommendation, where studies utilise AI to create effective learning groups; (2) group outcomes, where studies utilise AI to gather insights into the outcomes of learners' joint activities and (3) interaction analysis, where studies employ AI to examine the emotional conditions and tasks carried out by learners as they collaborate (Tan et al., 2022). In terms of data management opportunities promising directions include the use of reciprocal recommender systems that consider preferences of multiple learners (Palomares et al., 2021; Potts et al., 2018) in a privacy-preserving manner, and data modelling techniques that can measure team success, cohesion and types of interactions.

Teacher-supporting AI-EdTech

Growing attention is being given to the development of AI-EdTech aimed at supporting teachers. "Intelligent teaching assistant systems" aim to automate and extend the tasks that are usually completed by teaching assistants, such as grading (Yang & Zhu, 2022) and exam proctoring (Nigam et al., 2021). They can also be extended to support teachers in class orchestration

(Martinez-Maldonado, 2016), assessment (Swiecki et al., 2022), and detecting plagiarism (Foltýnek et al., 2019). Advancements in generative AI have also opened opportunities for instructors to partner with AI to "create educational content" such as learning objectives, lesson plans, learning resources and assessment items (Mhlanga, 2023). In addition, the wide availability of virtual and augmented reality technologies has resulted in the implementation of a wide range of "immersive learning environments" that are capable of simulating or imitating real and imagined worlds in safe and engaging environments. Despite advancements in AI-EdTech, instructors still encounter numerous obstacles when teaching, particularly in large classes. The utilisation of educational tools and the increasing occurrence of learning across physical and digital spaces make it arduous for instructors to obtain "actionable insights" (Jørnø & Gynther, 2018) into students' learning processes and progress to help them with their studies. A related challenge is that instructors find it hard to provide forms of guidance or "pedagogical interventions" that cater to the diverse academic abilities of learners. Finally, the continuous evolution of teaching tools and approaches necessitates conducting "empirical studies" to assess their effectiveness, which is often a challenging task.

Research and findings obtained from database and data management communities can be instrumental in resolving these issues. The use of online analytical processing (OLAP) techniques, including roll-up and drill-down and smart drill-downs can enable instructors to zoom in and out to explore student learning from different points of view (Shabaninejad et al., 2020), while applying descriptive, diagnostic, predictive, and prescriptive analytics can further facilitate generating actionable insights. The research and discoveries related to personalisation, recommender systems (Aggarwal et al., 2016), and nudging methods (Hummel & Maedche, 2019), alongside theoretical groundings from the learning sciences, can serve as the foundation for the creation of tailored pedagogical interventions aimed at supporting and enhancing student learning. Finally, to assess teaching tools and methods, data collection, storage, analysis, and sharing techniques utilised by the database and data management fields can aid in formulating performance metrics and creating scalable and reproducible experiments.

Institution-supporting AI-EdTech

In their systematic review of AI applications in higher education, Zawacki-Richter et al. (2019) noted that almost half (48%) of the included studies explored AI support for administrative and institutional services. Primary applications include "admission processes", such as making admission recommendations (Waters & Miikkulainen, 2014) with a focus on fairness and the institutions' reputations (Dennis, 2018; Marcinkowski et al., 2020; Zeide, 2019). There is also increasing use of university-wide intelligent "learning analytics dashboards" (LADs) that aim to help students, teachers and institutions understand and make informed decisions about learning processes (Khosravi et al., 2021; Matcha et al., 2019). Furthermore, there is a growing emphasis on the development of fair, accountable, and transparent "predictive learning analytics" that assist educators and administrative staff identify students who may be at risk of dropping out, failing a course, or experiencing poor well-being.

Although AI-EdTech offers advantages, its implementation also presents institutional challenges. Generating "data analytics" on student learning processes and experiences requires the consumption and integration of a variety of data sources, including student and curriculum records, learning logs and records, and data produced by educational tools, often in the face of questionable or at least unknown data quality. Further, the "ethics" of AI in education raises a variety of complex issues centred on data (e.g., consent and data privacy) and how that data is analysed (e.g., transparency and trust). The computational learner modelling employed by many AI-EdTech systems often uses profiles or stereotypes to predict academic performance and identify learners for early intervention (Chrysafiadi & Virvou, 2013). However, this approach can lead to discrimination in underrepresented populations (Sapiezynski et al., 2017). Inferring learner states from indicators or features such as gender, ethnic or cultural background, and socio-economic status also introduces bias and further widens existing gaps. Finally, to utilise effectively and to benefit from AI-EdTech, both students and educators need to possess adequate "data literacy" to access, interpret, and use data in educational contexts, which can be challenging to address on a large scale. Insights obtained from the database and data management communities can also play a key role in addressing these challenges. Data modelling and ML techniques can assist in developing explainable analytics (Shabaninejad et al., 2022), visualisations, and learner profiles (Barthakur et al., 2023) for comprehending the learning experiences and processes of students. Ethical use of data can be achieved by developing fair-by-design data access and processing primitives and fairness measures (Holstein & Doroudi, 2019) that protect minorities from being disadvantaged. Finally, partnering with experts in data and learning sciences provides opportunities for developing and delivering training workshops and learning modules on data literacy targeted towards students and instructors.

Conclusion and outlook

AI has the potential to address some of the biggest challenges in education today, advancing teaching and learning practices. In terms of supporting students, AI-powered technologies enable adaptive learning, collaborative learning and the development of various learner models and profiles that help students regulate and monitor their learning. In terms of supporting teachers AI-powered technologies aid teachers as virtual assistants in orchestrating classroom activities, facilitating immersive learning experiences, and automating tasks such as grading and responding to student queries. They can also help teachers in creating novel educational content as well as augmented and virtual learning experiences. At the institutional level, AI-powered technologies can help with admission process, generating dashboards, and by identifying dropouts and at-risk students.

While much has been achieved, data management challenges limit AI-EdTech from reaching its full potential. These challenges relate to capturing fine-grained data about learners and learning processes; including supporting learners' ability and agency to take ownership of their learning; incorporating in-database ML models to infer learners' data; developing algebraic primitives to express AI-EdTech applications; empowering teachers to make sense of student learning and have the technical capacity to intervene pedagogically to facilitate and provide tailored support; and ensuring that AI-EdTech is deployed with necessary ethical and responsible use guardrails.

We note the central role of data in education, and yet the enormous body of knowledge within the data management and data engineering community is largely absent from these discussions. In a recent keynote (Amer-Yahia, 2022) and publications (Khosravi et al., 2022; Swiecki et al., 2022), we flag this as a great opportunity for data management researchers to contribute to the development of AI-powered educational technologies by drawing upon their extant and emerging knowledge on the challenges outlined above; including large-scale and learning data-specific storage models; ability to reason and assess data properties and quality, real-time access and efficient indexing; powerful access primitives; query processing and inference engines; and supporting ethical and responsible use of data.

We thereby call upon educational experts and learning scientists to partner with the data management research community to overcome the plethora of data management challenges and thereby effectively respond to the changing global educational landscape.

Funding

This research was supported partially by the Australian Government through the Australian Research Council's Industrial Transformation Training Centre for Information Resilience (CIRES) project number IC200100022.

Disclosure statement

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this paper.

About the authors

Hassan Khosravi is currently an Associate Professor in Data Science and Learning Analytics at The University of Queensland. In his research, Hassan draws on theoretical insights from the learning sciences and exemplary techniques from the fields of human-computer interaction and human-centred AI to design, implement, validate, and deliver sociotechnical solutions that contribute to the delivery of learner-centred, data-informed learning at scale.

ORCID: https://orcid.org/0000-0001-8664-6117

Shazia Sadiq is a Professor in the School of Electrical Engineering and Computer Science at The University of Queensland, and Director for the ARC Centre for Information Resilience. Her research track record has focussed on overcoming challenges that stem from disparate IT systems and result in information silos, and she has developed new methods that to tackle these challenges through integrated solutions for information quality and effective use. She advocates for the responsible and ethical technology developments and believe strongly that these developments require trans-disciplinary collaborations between research, industry, government and community.

ORCID: https://orcid.org/0000-0001-6739-4145

Sihem Amer-Yahia is a Centre National de Recherche Scientifique (CNRS) Research Director and Deputy Director of the Laboratoire d'Informatique de Grenoble (LIG) in France. Her interests include the intersection of large-scale data management and social data exploration. She held positions at QCRI, Yahoo! Research and AT&T Labs. She served on the SIGMOD Executive Board, the VLDB Endowment, and the EDBT Board.

References

Abdi, S., Khosravi, H., & Sadiq, S. (2021). Modelling learners in adaptive educational systems: A multivariate glicko-based approach. In M. Scheffel, N. Dowell, S. Joksimović, & G. Siemens (Eds.), LAK '21: *Proceedings of the 11th International Conference on Learning Analytics & Knowledge* (pp. 497–503). Association for Computing Machinery.

Abdi, S., Khosravi, H., Sadiq, S., & Gašević, D. (2020). Complementing educational recommender systems with open learner models. In V. Kovanović, M. Scheffel, N. Pinkwart, & K. Verbert (Eds.), *LAK '20: Proceedings of the 10th International Conference on Learning Analytics & Knowledge* (pp. 360–365). Association for Computing Machinery.

Aggarwal, C. C. et al. (2016). *Recommender systems* (Vol. 1). Springer.

Amer-Yahia, S. (2022). Towards AI-powered data-driven education. *Proceedings of the VLDB Endowment*, *15*(12), 3798-3806. https://doi.org/10.14778/3554821.3554900

Amer-Yahia, S., & Roy, S. B. (2016). Human factors in crowdsourcing. *Proceedings of the VLDB Endowment, 9*(13), 1615–1618. https://doi.org/10.14778/3007263.3007323

Azzalini, D., Azzalini, F., Criscuolo, C., Dolci, T., Martinenghi, D., & Amer-Yahia, S. (2022). SoCRATe: A recommendation system with limited-availability items. In *CIKM '22: Proceedings of the 31st ACM International Conference on Information & Knowledge Management* (pp. 4793–4797). ACM. https://doi.org/10.1145/3511808.3557208

Barthakur, A., Dawson, S., & Kovanovic, V. (2023). Advancing leaner profiles with learning analytics: A scoping review of current trends and challenges. In I. Hilliger, H. Khosravi, B. Rienties, & S. Dawson (Eds.), *LAK '23: Proceedings of the 13th International Conference on Learning Analytics & Knowledge* (pp. 606–612). Association for Computing Machinery.

Becker, B. (2017). Artificial intelligence in education: What is it, where is it now, where is it going. *Ireland's Yearbook of Education*, 2018, 42-46.

Briggs, K. C. (1976). *Myers-Briggs type indicator*. Consulting Psychologists Press.

Bull, S. (2020). There are open learner models about! *IEEE Transactions on Learning Technologies*, *13*(2), 425–448. https://doi.org/10.1109/TLT.2020.2978473

Cao, H., Yang, V., Chen, V., Lee, Y. J., Stone, L., Diarrassouba, N. J., Whiting, M. E., & Bernstein, M. S. (2020). My team will go on: Differentiating high and low viability teams through team interaction. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW3), Article 230, 1–27. https://doi.org/10.48550/arXiv.2010.07292

Chrysafiadi, K., & Virvou, M. (2013). Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, *40*(11), 4715–4729. https://doi.org/10.1016/j.eswa.2013.02.007

Darvishi, A., Khosravi, H., Sadiq, S., & Weber, B. (2022). Neurophysiological measurements in higher education: A systematic literature review. *International Journal of Artificial Intelligence in Education*, *32*(2), 413–453. https://doi.org/10.1007/s40593-021-00256-0

Dennis, M. J. (2018). Artificial intelligence and recruitment, admission, progression, and retention. *Enrollment Management Report*, *22*(9), 1-3. https://doi.org/10.1002/emt.30479

Fadel, C., Holmes, W., & Bialik, M. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. The Center for Curriculum Redesign, Boston, MA.

Foltýnek, T., Meuschke, N., & Gipp, B. (2019). Academic plagiarism detection: A systematic literature review. *ACM Computing Surveys*, *52*(6), 1–42. https://doi.org/10.1145/3345317

Ge, Y., Liu, S., Gao, R., Xian, Y., Li, Y., Zhao, X., Pei, C., Sun, F., Ge, J., Ou, W., & Zhang, Y. (2021). Towards long-term fairness in recommendation. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining* (pp. 445–453). Association for Computing Machinery. https://doi.org/10.48550/arXiv.2101.03584

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable artificial intelligence. *Science Robotics*, *4*(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Hackman, J., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, *16*(2), 250-279. https://doi.org/10.1016/0030-5073(76)90016-7

Hata, K., Krishna, R., Fei-Fei, L., & Bernstein, M. S. (2017). A glimpse far into the future: Understanding long-term crowd worker quality. In C. P. Lee, S. Poltrock, L. Barkhuus, M. Borges & W. Kellogg (Eds.), *CSCW '17: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 889–901). Association for Computing Machinery. https://doi.org/10.48550/arXiv.1609.04855

Holstein, K., & Doroudi, S. (2019). Fairness and equity in learning analytics systems (FairLAK). In J. Cunningham, N. Hoover, S. Hsiao, G. Lynch, K. McCarthy, C. Brooks, R. Ferguson, & U. Hoppe (Eds.), *Companion proceedings 9th International Conference on Learning Analytics & Knowledge (LAK19)* (pp. 500–503). Society for Learning Analytics Research.

Hummel, D., & Maedche, A. (2019). How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics*, *80*, 47–58. https://doi.org/10.1016/j.socec.2019.03.005

Jørnø, R. L., & Gynther, K. (2018). What constitutes an "actionable insight" in learning analytics? *Journal of Learning Analytics*, *5*(3),198–221. https://doi.org/10.18608/jla.2018.53.13

Kaufmann, N., Schulze, T., & Veit, D. (2011). More than fun and money. Worker motivation in crowdsourcing – A study on mechanical turk. In V. Sambamurthy, & M. Tanniru (Eds.), *A renaissance of information technology for sustainability and global competitiveness. 17th Americas Conference on Information Systems, AMCIS 2011*. Association for Information Systems.

Kizilcec, R. F., & Lee, H. (2020). Algorithmic fairness in education. *arXiv:2007.05443*. https://doi.org/10.48550/arXiv.2007.05443

Khosravi, H., Shabaninejad, S., Bakharia, A., Sadiq, S., Indulska, M., & Gašević, D. (2021). Intelligent learning analytics dashboards: Automated drill-down recommendations to support teacher data exploration. *Journal of Learning Analytics*, *8*(3), 133–154. https://doi.org/10.18608/jla.2021.7279

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *3*,100074. https://doi.org/10.1016/j.caeai.2022.100074

Luo, S., Gao, Z. J., Gubanov, M. N., Perez, L. L., Jankov, D., & Jermaine, C. M. (2020). Scalable linear algebra on a relational database system. *Communications of the ACM*, *63*(8), 93–101. https://doi.org/10.1109/ICDE.2017.108

Marcadet, G., Ciucanu, R., Lafourcade, P., Soare, M., & Amer-Yahia, S. (2022). Samba: A system for secure federated multi-armed bandits. In *2022 IEEE: 38th International Conference on Data Engineering (ICDE)* (pp. 3154–3157). IEEE. https://doi.org/10.1109/ICDE53745.2022.00286

Marcinkowski, F., Kieslich, K., Starke, C., & Lünich, M. (2020). Implications of AI (un-)fairness in higher education admissions: The effects of perceived AI (un-)fairness on exit, voice and organizational reputation. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 122–130). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372867

Martinez-Maldonado, R. (2016). Seeing learning analytics tools as orchestration technologies: Towards supporting learning activities across physical and digital spaces. *CEUR Workshop Proceedings, 1601*, 70–73.

Matcha, W., Gašević, D., & Pardo, A., et al. (2019). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226–245. https://doi.org/10.1109/TLT.2019.2916802

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. *SSRN Electronic Journal.* https://doi.org/10.2139/SSRN.4354422

Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policymakers*. UNESCO Publishing.

Nigam, A., Pasricha, R., Singh, T., & Churi, P. (2021). A systematic review on AI-based proctoring systems: Past, present and future. *Education and Information Technologies*, *26*(5), 6421–6445. https://doi.org/10.1007/s10639-021-10597-x

Palomares, I., Porcel, C., Pizzato, L., Guy, I., & Herrera-Viedma, E. (2021). Reciprocal recommender systems: Analysis of state-of-art literature, challenges and opportunities towards social recommendation. *Information Fusion*, *69*, 103–127. https://doi.org/10.1016/j.inffus.2020.12.00

Pilourdault, J., Amer-Yahia, S., Lee, D., & Roy, S. B. (2017). Motivation-aware task assignment in crowdsourcing. In V. Markl, S. Orlando, B. Mitschang, P. Andritsos, K.-U. Sattler, & S. Breß (Eds.), *Proceedings of the 20th International Conference on Extending Database Technology, EDBT 2017* (pp. 246–257). EDBT. https://doi.org/10.5441/002/edbt.2017.23

Pilourdault, J., Amer-Yahia, S., Roy, S. B., & Lee, D. (2018). Task relevance and diversity as worker motivation in crowd-sourcing. In *34th IEEE International Conference on Data Engineering, ICDE 2018* (pp. 365–376). IEEE. https://doi.org/10.1109/ICDE.2018.00041

Posch, L., Bleier, A., Lechner, C., Danner, D., Flöck, F., & Strohmaier, M. (2019). Measuring motivations of crowdworkers: The multidimensional crowdworker motivation scale. *ACM Transactions on Social Computing*, *2(*2), 8:1–8:34. https://doi.org/10.1145/3335081

Potts, B. A., Khosravi, H., Reidsema, C., Bakharia, A., Belonogoff, M., & Fleming, M. (2018). Reciprocal peer recommendation for learning purposes. In A. Pardo, K. Bartimote-Aufflick, G. Lynch, S. Buckingham Shum, R. Ferguson, A. Merceron, & X. Ochoa (Eds.), *LAK '18: Proceedings of the 8th International Conference on Learning Analytics & Knowledge* (pp. 226–235). Association for Computing Machinery. https://doi.org/10.1145/3170358.3170400

Rahman, H., Thirumuruganathan, S., Roy, S. B., Amer-Yahia, S., & Das, G. (2015). Worker skill estimation in team-based tasks. *Proceedings of the VLDB Endowment, 8*(11), 1142–1153. https://doi.org/10.14778/2809974.2809977

Rastrollo-Guerrero, J. L., Gómez-Pulido, J. A., & Durán-Domínguez, A. (2020). Analyzing and predicting students' performance by means of machine learning: A review. *Applied Sciences*, *10*(3), 1042.

Roy, S. B., Lykourentzou, I., Thirumuruganathan, S., Amer-Yahia, S., & Das, G. (2015). Task assignment optimization in knowledge-intensive crowdsourcing. *The VLDB Journal*, *24*(4), 467–491. https://doi.org/10.1007/s00778-015-0385-2

Sapci, A. H., & Sapci, H. A. (2020). Artificial intelligence education and tools for medical and health informatics students: Systematic review. *JMIR Medical Education*, *6*(1), e19285. https://doi.org/10.2196/19285

Sapiezynski, P., Kassarnig, V., & Wilson, C. (2017). *Academic performance prediction in a gender-imbalanced environment* [Data set]. Boise State University, ID.

Shabaninejad, S., Khosravi, H., Abdi, S., Indulska, M., & Sadiq, S. (2022). Incorporating explainable learning analytics to assist educators with identifying students in need of attention. In *Proceedings of the Ninth ACM Conference on Learning@Scale* (pp. 384–388). Association for Computing Machinery. https://doi.org/10.1145/3491140.3528292

Shabaninejad, S., Khosravi, H., Indulska, M., Bakharia, A., & Isaias, P. (2020). Automated insightful drill-down recommendations for learning analytics dashboards. In V. Kovanović, M. Scheffel, N. Pinkwart, & K. Verbert (Eds.), *LAK '20: Proceedings of the 10th International Conference on Learning Analytics*

& Knowledge (pp. 41–46). Association for Computing Machinery. https://doi.org/10.1145/3375462.3375539

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, *3*, 100075. https://doi.org/10.1016/j.caeai.2022.100075

Tan, S. C., Lee, A. V. Y., & Lee, M. (2022). A systematic review of artificial intelligence techniques for collaborative learning over the past two decades. *Computers and Education: Artificial Intelligence*, 100097. https://doi.org/10.1016/j.caeai.2022.100097

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221. https://doi.org/10.1080/00461520.2011.611369

Waters, A., & Miikkulainen, R. (2014). GRADE: Machine learning support for graduate admissions. *AI Magazine*, *35*(1), 64–75. https://doi.org/10.1609/aimag.v35i1.2504

Yang, H., & Zhu, K. (2022). Research on AI-assisted grading of math questions based on deep learning. In P. Kar; & S. Guan (Eds.), *International Conference on Internet of Things and Machine Learning (IoTML 2021)* (vol. 12174), (pp. 211–217). SPIE. https://doi.org/10.1117/12.2628649

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

Zeide, E. (2019). Artificial intelligence in higher education: Applications, promise and perils, and ethical questions. *Educause Review*, *54*(3).