A Computing Education Perspective on AI Explainability

Lana Issa[0000−0002−9440−3366] and Ralf Romeike[0000−0002−2941−4288]

Freie Universität Berlin, Berlin, Germany
{lana.issa, ralf.romeike}@fu-berlin.de

Abstract. The integration of artificial intelligence (AI) systems into educational environments presents both opportunities and challenges for teaching and learning. Although ethical frameworks emphasise principles such as fairness, accountability, and transparency, their operationalisation in educational AI remains difficult, particularly with respect to explainability. This paper approaches explainability from a computing education perspective, arguing that computing educators play a key role in supporting transparency by fostering user understandability through pedagogical practice. We bring together ethical concerns about AI in education and create a framework for user-centred explainability that is grounded in theory, with pedagogically structured explanations. To explain this framework in action, we present a small-scale case study as an illustrative instantiation of how it works within an AI-based educational tool. The paper contributes a conceptual foundation for positioning computing educators as key actors in strengthening understandability and trust in educational AI systems.

Keywords: AI explainability · AI transparency · Computing Education

1 Introduction

Artificial Intelligence (AI) technologies have seen rapid expansion in many domains, including business, healthcare, and increasingly education [19]. Modern AI systems often consist of multiple interdependent modules trained on large-scale datasets and incorporating advanced techniques such as deep learning [5]. While these developments offer considerable potential, they also introduce challenges concerning understanding of how these systems operate and how they generate output [14]. In response to these concerns, international guidelines and evaluation frameworks increasingly emphasise ethical principles such as fairness, accountability, privacy, and transparency as foundations for trustworthy AI [4,10,36]. While these guidelines establish a valuable theoretical foundation for AI ethics, there is still a need for practical frameworks and taxonomies that articulate the concrete practices required to achieve trustworthy AI and adapt them to specific contexts [4], because for non-expert users, it is often unclear how decisions are made, what data the system was trained on, and what limitations, assumptions, or uncertainties shape its behaviour, raising important questions about when and how such systems can be responsibly used [28].

Transparency involves making visible the system's inner components, data flows, assumptions, and ethical considerations [22]. One practical dimension of transparency is explainability, which refers to communicating an AI system's capabilities, decision logic, and outputs in ways that allow users to grasp why a particular result was produced [18]. Because users in different domains have different goals, levels of expertise, and explanatory needs, it is challenging to limit explainability to a single definition or method [3]. One effective way to evaluate explanations is to assess how well they are interpreted by users, since different settings require different explanations and users need to understand AI systems' inner workings sufficiently to interpret and question their outputs [25]. This has motivated research on creating AI explainability techniques that are interpretable [13,20], yet a gap persists between technical explanations and end-user understanding, often reflected in uncertainty or insufficient trust [39]. Consequently, there is a growing emphasis on understandability, which we define as the user's ability to form accurate, meaningful mental models of an AI system's behaviour, data use, limitations, and implications in their specific context, supported through pedagogically structured explanations [12]. Approaches grounded in AI literacy, including the Dagstuhl AI literacy triangle [31], highlight the need for conceptual, socio-ethical, and practical understanding of AI technologies, reinforcing the importance of pedagogical explanations, accessible representations, and plain-language descriptions of system behaviour [9]. Such understandability fosters real trust rather than blind confidence and enables more informed use of AI systems [11].

Building on this perspective, this paper argues that computing education can play a mediating role in AI interpretability by foregrounding understandability as a pedagogical construct. Rather than focusing on model-level explainability as studied in Explainable Artificial Intelligence (XAI), this work addresses explainability as a user-facing, pedagogical process aimed at fostering understandability. We conceptualise understandability as the outcome of pedagogically designed explanations that enable users to make sense of AI systems in context. Computing educators have the ability to translate complex model behaviour, data processes, and system limitations into clear and learnable representations that fosters understandability for a diversity of users. In this paper, we create a theoretical framework for embedding understandability into AI systems, grounded in principles of AI ethics and educational thinking. Using the domain of education as an example, we first express key ethical considerations that affects users' understandability requirements. We then introduce four practical concepts: AI fundamentals, a modular "glass-box" system map, an overview of data life-cycle, and an ethics and limitations card. These points form aconceptual framework for operationalising understandability in actual settings. Then, we present a small exploratory case study, in which these concepts were applied within an AI-based reflective writing tool, mainly to reflect on our own framework instead of only evaluating results. We conclude by deriving general observations within the field of education to highlight the particular ethical considerations of audience in the domain of education and outline how computing educators can systematically support AI understandability beyond this specific case.

2 Prior Literature

Although discussions about explainability are increasing, detailed practices in specific domains are still needed [4], and more comprehensive work is required on AI acceptance and trustworthiness across contexts [29]. Several researchers have analysed general AI ethics principles within specific domains to establish foundations for more applicable and tailored explainability. For example, Chaudhry et al. [7] proposed a transparency index for educational AI, Holmes et al. [19] developed a framework addressing ethical concerns around educational data, and Nguyen et al. [35] examined policies and regulations related to AI ethics in education. These efforts underscore the need to translate ethical guidelines into domain-relevant explainability practices that enhance trustworthiness [6,30].

Research on AI literacy provides a basis for understanding how users develop the conceptual, socio-ethical, and practical competencies necessary to make sense of AI systems. The Dagstuhl AI Literacy Framework situates AI understanding across these three areas [31], emphasizing that interpretability is both technical and educational. Users require structured, pedagogical explanations to grasp data flows, model logic, and system limitations [26], aligning with broader arguments that interpretability emerges when explanations connect system behaviour to underlying concepts rather than merely revealing technical mechanisms [12].

Work on explainability in educational and user-centred AI highlights the persistent gap between technical XAI outputs and the explanations non-expert users actually need. Reviews show that mainstream explainability tools often do not support meaningful understanding [33], while educational studies demonstrate that contextualised and cognitively manageable explanations aligned with user goals improve comprehension [37]. Empirical research illustrates that structured documentation, modular breakdowns, and visual data-flow representations support users in forming actionable mental models of AI behaviour [27,42]. Complementary work on transparency and responsible design reinforces the value of accessible documentation and socio-ethical framing [4,16,32]. Studies in computing education further show how educators can translate complex AI components into learnable concepts that scaffold responsible engagement with AI tools [40]. Collectively, these findings indicate that computing educators can serve as mediators of AI interpretability by designing explanations that enhance users' understandability across domains.

3 Methodology

This study adopts a theory-driven design approach to articulate how understandability can be systematically embedded into AI systems through pedagogically grounded explanatory materials aligned with established ethical AI frameworks.

The main objective is not to assess one specific system, but to develop and illustrate a conceptual framework to explainability that is focused on the audience, connects ethical principles, educational goals, and technical characteristics of systems. The framework is grounded in the context of a university course on digital and AI literacy, where students interact with an AI-based reflective writing tool, providing a realistic setting to illustrate how such understandability-oriented explanations can support user sensemaking, transparency, and informed trust.

3.1 Theoretical Grounding

To ground the design of the explanatory materials in recognised ethical and educational standards, we conducted a review of three foundational bodies of work:

- **European Commission Ethics Guidelines and regulatory documents** [10,36]: These provide the core qualities of "Trustworthy AI", including transparency, accountability, fairness, safety, robustness, and human oversight.
- **German Standardisation Roadmap for AI** [28]: This document complements the European guidelines with operationalisation-oriented qualities such as traceability, explainability, reproducibility, and documentation practices relevant for deployment in public and educational institutions.
- **Dagstuhl AI Literacy Framework** [31]: This framework contributes user-centred insights into what learners need to understand about AI in terms of concepts, applications, and societal perspective, thus helping translate high-level ethical principles into educational requirements.

Table 1: Overlapping and distinct trustworthy AI requirements in the EU and German roadmap.

Theme	**EU Trustworthy AI**	**German AI Roadmap**
Transparency and Explainability	Clear purpose, logic, limits; context-appropriate explanations.	Documentation standards; traceability and interpretable modular design.
Human Oversight	Meaningful control, prevention of automation bias.	Procedural oversight and governance structures.
Accountability	Responsibility allocation; mechanisms for redress.	Auditability, technical governance, risk management.
Fairness	Avoid unjust bias and ensure equitable treatment.	Bias testing, dataset documentation, mitigation workflows.
Privacy and Data Governance	Lawful, secure, and ethical data handling.	Operational data governance and traceable data life cycles.
Robustness and Safety	Reliable and resilient system behaviour.	Verification, validation, reproducibility standards.

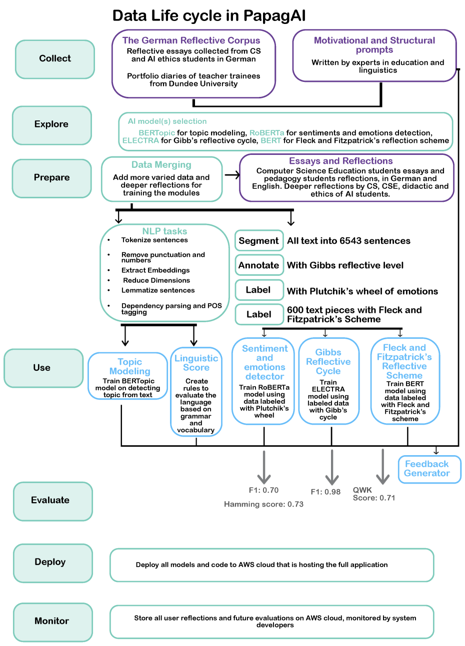
Based on these sources, we identified a list of transparency and explainability requirements that focus on supporting user understandability, to inform our design of the explanatory materials. Table 1 shows the list of requirements.

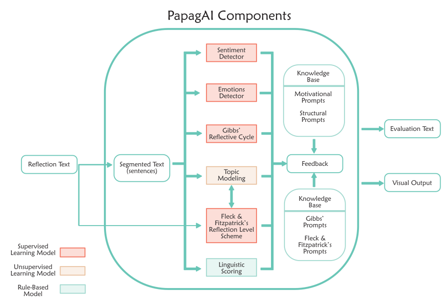
3.2 Conceptual Framework for Audience-centred Understandability

We develop four explanatory components to serve as a framework for operationalising understandability in AI-based educational systems. Rather than viewing explanations as isolated features, this framework considers explainability as a pedagogical scaffold that connects ethical AI principles, educational goals, and technical system characteristics. Each component addresses a distinct dimension of user sensemaking and together they provide a structured approach for translating abstract ethical requirements into concrete, learnable representations aligned with system functionality.

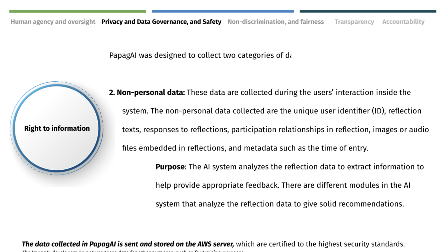
– **AI Fundamentals Micro-Units (Conceptual Grounding):** These micro-units provide learners with concise, modular introductions to the specific AI techniques employed by the system (e.g., machine learning, sentiment analysis, and large language models). The micro-units are directly linked to the system's decision points, enabling users to link observed outputs to underlying computational processes rather than providing AI literacy content. This component satisfies the ethical requirement of transparency by supporting the formation of accurate mental models and lowering the possibility of over- or under-trusting the system's feedback.
– **Modular "Glass-Box" System Map (Structural Transparency):** The system map externalises the inner structure of the AI application by decomposing it into functionally meaningful modules. It illustrates how input text is processed throughout multiple AI and non-AI components and how every partial result contributes to the final feedback score. By highlighting the relationships between modules, this component provides a view towards structural transparency, supporting traceability, accountability, and informed questioning of system behaviour.
– **Data Life-Cycle Overview (Data-centred Understandability):** This component foregrounds the imperative role of data in shaping AI behaviour by making the entire data life cycle visible to users, informed by the principles articulated by Olari and Romeike [38]. It explains how user inputs are collected, processed, and moved through the system; it differentiates between training data and live data, and communicates the accuracy of individual AI modules that manipulate these data. By linking data sources, processing steps, and model performance to outputs, this overview supports ethical principles of fairness, data governance, and responsible reliance on AI-generated feedback.
– **Ethics Awareness Cards (User Agency)**: The ethics cards translate abstract principles of AI ethics into accessible guidance for users. They summarise the system's purpose, capabilities, limitations, privacy considerations, data governance and accountability practices. This component strengthens

human oversight by clearly stating what the system can and cannot do, outlining user responsibilities, and encouraging reflective and critical engagement rather than passive acceptance of AI outputs.





(b) PapagAI Modular Components.



(a) Data Lifecycle in PapagAI System.

(c) A sample Card of the Card-Based Ethical Report of PapagAI.

Fig. 1: Visual elements of the PapagAI system, including data lifecycle, system components, and ethical explanations.

Together, these four components constitute a pedagogical *understandability framework* that embeds ethical AI principles into the concrete operation of an educational AI system. Rather than serving as isolated explanatory features, they function collectively as an audience-centred explainability layer that enables users to interpret, question, and appropriately trust AI-supported feedback. Figure 1 presents illustrative excerpts from these materials.

3.3 Framework Application in an Educational Context

To illustrate the proposed framework, we applied it within a university-wide course on digital and AI literacy, enrolling approximately 200 undergraduate students from different disciplines. As part of the course, all students used PapagAI [45], an AI-based reflective writing tool offering automated feedback on linguistic clarity, reflective depth, emotional tone, and writing quality.

The study included a pre-survey (administered before exposure to the explanatory materials) and a post-survey (administered after using the tool and engaging with the materials), the surveys aim at capturing the perceived clarity of explanations, improved understanding of system behaviour, and changes in trust or confidence. The surveys included Likert-scale questions. Although the sample size (n = 18) limits generalisability, it offers preliminary insight into whether pedagogically structured explanatory materials can enhance users' understandability and system acceptance.

4 Results and Discussion

4.1 Survey results

The main goal of the survey was to illustrate the proposed understandability framework and gather initial indications of its effectiveness. Specifically, it aimed to assess whether the provided explanatory materials enhanced students' perception of transparency, supported trust in the AI system, and improved their understanding of how the system operates. Table 2 presents the survey results.

Table 2: Student survey results

Category	Survey Statement	Pre (%)	Post (%)	Change
General	I would use such an AI application to reflect on my classroom experience	45.5	66.7	+21.2
	The topic of AI should be covered in more studies	64.1	80.0	+15.9
Attitude, bias, and fear	I am concerned that the AI system could limit or misdirect humans' autonomy	60.0	38.9	–21.1
	I fear that I may communicate with the AI system without realizing it	65.1	41.1	–24.0
	I fear AI systems will reinforce stereotypes and increase discrimination	53.7	61.1	+7.4
Ethical and legal concerns	I fear that AI systems could violate privacy and data protection rights	60.0	75.0	+15.0
	I'm afraid AI systems could cause harm due to technical problems	64.9	66.7	+1.8
Motivation	Would you use AI-supported systems to help with career decisions	39.8	55.5	+15.7
	I assume AI knowledge will be useful in my future profession	12.1	40.0	+27.9

The pre- and post-survey results indicate a general improvement in student attitudes toward the use of AI in education following their exposure to the explanatory materials. In the general category, the numbers suggest increased interest and perceived relevance of AI in educational contexts.

In the category of attitudes, bias, and fear, concerns about AI undermining human autonomy and interacting with AI systems unknowingly showed notable decreases (–21.1% and –24%, respectively), indicating reduced fear and improved trust after system use. However, fear related to AI that reinforces social biases increased slightly (+7.4%), which may reflect a greater critical awareness of ethical implications rather than a negative change in perception.

Under ethical and legal concerns, fear of privacy violations increased by 15%, and concern over technical errors causing harm increased slightly by 1.8%. This rise may suggest that exposure to the system, while reducing irrational fears, also brought legitimate ethical issues more clearly into focus, possibly due to the transparency mechanisms making risks more visible and discussable.

Lastly, students showed substantial growth in both AI acceptance and future-oriented interest. Positive responses to using AI for career guidance rose by 15.7%, while the belief in AI's professional utility jumped by 27.9%, the largest observed increase in the survey. This suggests that explainability not only fosters trust, but also positively affects learners' motivations to engage with AI systems in their academic and professional development.

4.2 Ethical Considerations for AI in Educational Contexts

The survey findings align with prior work showing that, although ethical principles such as fairness, transparency, and accountability are widely acknowledged in AI research [14], their practical realisation in education continues to present difficulties. Our results confirm the concerns raised in the literature [19,34] that educational contexts require domain-tailored transparency implementations, particularly because AI systems influence learning, assessment, and student support. The observed user perceptions reinforce the argument that ethical principles must be paired with concrete, audience-oriented explanations to become meaningful for learners and educators.

Ethical and legal concerns in educational data
Participants' concerns around data use and privacy reflect the issue in using AI systems in education, where personal and sensitive learner data must be managed under strict legal and ethical safeguards [1,24]. Consistent with General Data Protection Regulation (GDPR) obligations [47] and broader international frameworks [10,28,36], students expressed a desire for clarity on what data are collected, how they are processed, and for what purposes. This aligns with long-standing work in learning analytics emphasizing the centrality of transparent data governance in fostering trust [15,44]. Our findings suggest that explanatory materials that make data life cycles visible and understandable directly address these expectations.

Attitudes, biases, and perceived risks of educators and learners
Survey results further resonate with research indicating that fears of bias and uncertainty about system reliability shape user acceptance in education [41]. Prior work suggests that transparency about model behaviour, development practices, and limitations can mitigate such concerns [2,9,17]. Our explanatory materials aimed to surface precisely these elements, and the reported increase in perceived system clarity supports arguments that audience-centred explainability fosters

trust. The results also support findings that social and institutional transparency contributes to greater confidence in AI-based systems in educational contexts [14].

Cognitive, emotional, motivational, and volitional conditions
Opaque AI systems impose unnecessary interpretive burden on learners [7,8,23]. Such burdens can generate frustration and interfere with the learning process [46]. Conversely, well-designed explanations and interfaces can support motivation, autonomy, and comfort, which are key factors emphasised in human-centred educational technology design [21,43]. The improved perceptions of system understandability observed in our study suggest that tailored explanatory materials can meaningfully reduce cognitive strain and enhance users' willingness to engage with AI-based tools.

5 Conclusion and Future Work

This paper presents a theoretical perspective on how computing education can contribute to transparency and trust in AI-based educational systems. By integrating measures from AI ethics, explainability, and concepts in education, we developed a framework to support audience-centred understandability. The presented explanatory materials demonstrate how ethical principles can be operationalised into the design and components of AI systems, with a case study serving to illustrate their application. Our contribution emphasises the role of computing educators as mediators of AI transparency rather than evaluators of specific tools. For future work, we encourage educators to apply and test the framework in their own classrooms and with different AI systems to explore its practical impact and transferability.

References

1. Amo, D., Torres, R., Canaleta, X., Herrero-Martín, J., Rodríguez-Merino, C., Fonseca, D.: Seven principles to foster privacy and security in educational tools: Local educational data analytics. In: Eighth international conference on technological ecosystems for enhancing multiculturality. pp. 730–737 (2020)
2. Anik, A.I., Bunt, A.: Data-centric explanations: explaining training data of machine learning systems to promote transparency. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. pp. 1–13 (2021)
3. Arya, V., Bellamy, R.K.E., Chen, P., Dhurandhar, A., Hind, M., Hoffman, S.C., Houde, S., Liao, Q.V., Luss, R., Mojsilovic, A., Mourad, S., Pedemonte, P., Raghavendra, R., Richards, J.T., Sattigeri, P., Shanmugam, K., Singh, M., Varshney, K.R., Wei, D., Zhang, Y.: One explanation does not fit all: A toolkit and taxonomy of AI explainability techniques. CoRR **abs/1909.03012** (2019), http://arxiv.org/abs/1909.03012
4. Balasubramaniam, N., Kauppinen, M., Hiekkanen, K., Kujala, S.: Transparency and explainability of ai systems: ethical guidelines in practice. In: International Working Conference on Requirements Engineering: Foundation for Software Quality. pp. 3–18. Springer (2022)

5. Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., et al.: On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258 (2021), https://arxiv.org/abs/2108.07258
6. Borenstein, J., Howard, A.: Emerging challenges in ai and the need for ai ethics education. AI and Ethics **1**, 61–65 (2021)
7. Chaudhry, M.A., Cukurova, M., Luckin, R.: A transparency index framework for ai in education. In: International Conference on Artificial Intelligence in Education. pp. 195–198. Springer (2022)
8. Chazette, L., Karras, O., Schneider, K.: Do end-users want explanations? analyzing the role of explainability as an emerging aspect of non-functional requirements. In: 2019 IEEE 27th international requirements engineering conference (RE). pp. 223–233. IEEE (2019)
9. Chikobava, M., Romeike, R.: Towards an operationalization of ai acceptance among pre-service teachers. In: Proceedings of the 16th Workshop in Primary and Secondary Computing Education. pp. 1–2 (2021)
10. Commission, E., for Communications Networks, D.G., Content, Technology: Ethics guidelines for trustworthy AI. Publications Office (2019). https://doi.org/doi/10.2759/346720
11. Cysneiros, L.M., Raffi, M., do Prado Leite, J.C.S.: Software transparency as a key requirement for self-driving cars. In: 2018 IEEE 26th international requirements engineering conference (RE). pp. 382–387. IEEE (2018)
12. Erasmus, A., Brunet, T.D., Fisher, E.: What is interpretability? Philosophy & Technology **34**(4), 833–862 (2021)
13. Feldhus, N., Kopf, L.: Interpreting language models through concept descriptions: A survey. In: Proceedings of the 8th BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP. pp. 149–162 (2025)
14. Felzmann, H., Villaronga, E.F., Lutz, C., Tamò-Larrieux, A.: Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. Big Data & Society **6**(1), 2053951719860542 (2019)
15. Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T., Vuorikari, R.: Research evidence on the use of learning analytics: Implications for education policy (2016)
16. Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J.W., Wallach, H., Iii, H.D., Crawford, K.: Datasheets for datasets. Communications of the ACM **64**(12), 86–92 (2021)
17. Grande, V., Kinnunen, P., Peters, A.K., Barr, M., Cajander, r., Daniels, M., Lewis, A.N., Sabin, M., Sánchez-Peña, M., Thota, N.: Role modeling as a computing educator in higher education: A focus on care, emotions and professional competencies. Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3571785.3574122, https://doi.org/10.1145/3571785.3574122
18. Habibullah, K.M., Horkoff, J.: Non-functional requirements for machine learning: Understanding current use and challenges in industry. In: 2021 IEEE 29th International Requirements Engineering Conference (RE). pp. 13–23. IEEE (2021)
19. Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S.B., Santos, O.C., Rodrigo, M.T., Cukurova, M., Bittencourt, I.I., et al.: Ethics of ai in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education pp. 1–23 (2022)
20. Holzinger, A., Saranti, A., Molnar, C., Biecek, P., Samek, W.: Explainable ai methods-a brief overview. In: International workshop on extending explainable AI beyond deep models and classifiers. pp. 13–38. Springer (2022)

21. Issa, L., Jusoh, S.: Usability evaluation on gamified e-learning platforms. In: Proceedings of the Second International Conference on Data Science, E-Learning and Information Systems. pp. 1–6 (2019)
22. Karran, A.J., Charland, P., Trempe-Martineau, J., Ortiz de Guinea Lopez de Arana, A., Lesage, A.M., Senecal, S., Leger, P.M.: Multi-stakeholder perspective on responsible artificial intelligence and acceptability in education. npj Science of Learning **10**(1), 44 (2025)
23. Köhl, M.A., Baum, K., Langer, M., Oster, D., Speith, T., Bohlender, D.: Explainability as a non-functional requirement. In: 2019 IEEE 27th International Requirements Engineering Conference (RE). pp. 363–368. IEEE (2019)
24. Kronivets, T., Yakovenko, O., Tymoshenko, Y., Ilnytskyi, M., Iasechko, S., Iasechko, M., et al.: Legal and ethical dimensions of ai in education: Navigating new frontiers. Review of Artificial Intelligence in Education **4**, e21–e21 (2023)
25. Lan, Y.J., Chen, N.S.: Teachers' agency in the era of llm and generative ai. Educational Technology & Society **27**(1), I–XVIII (2024)
26. Laupichler, M.C., Aster, A., Schirch, J., Raupach, T.: Artificial intelligence literacy in higher and adult education: A scoping literature review. Computers and Education: Artificial Intelligence **3**, 100101 (2022)
27. Lee, M.H., Choo, S.X.Y., Thilarajah, S.D., et al.: Improving health professionals' onboarding with ai and xai for trustworthy human-ai collaborative decision making. arXiv preprint arXiv:2405.16424 (2024)
28. Lepri, B., Oliver, N., Letouzé, E., Pentland, A., Vinck, P.: Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. Philosophy & Technology **31**(4), 611–627 (2018)
29. Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., Yi, J., Zhou, B.: Trustworthy ai: From principles to practices. ACM Computing Surveys **55**(9), 1–46 (2023)
30. Miao, F., Holmes, W., Huang, R., Zhang, H., et al.: AI and education: A guidance for policymakers. UNESCO Publishing (2021)
31. Michaeli, T., Romeike, R., Seegerer, S.: What students can learn about artificial intelligence–recommendations for k-12 computing education. In: IFIP World Conference on Computers in Education. pp. 196–208. Springer (2022)
32. Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I.D., Gebru, T.: Model cards for model reporting. In: Proceedings of the conference on fairness, accountability, and transparency. pp. 220–229 (2019)
33. Mohseni, S., Zarei, N., Ragan, E.D.: A multidisciplinary survey and framework for design and evaluation of explainable ai systems. ACM Transactions on Interactive Intelligent Systems (TiiS) **11**(3-4), 1–45 (2021)
34. Ng, D.T.K., Leung, J.K.L., Su, J., Ng, R.C.W., Chu, S.K.W.: Teachers' ai digital competencies and twenty-first century skills in the post-pandemic world. Educational technology research and development **71**(1), 137–161 (2023)
35. Nguyen, A., Ngo, H.N., Hong, Y., Dang, B., Nguyen, B.P.T.: Ethical principles for artificial intelligence in education. Education and Information Technologies **28**(4), 4221–4241 (2023)
36. Nikolinakos, N.T.: A european approach to excellence and trust: The 2020 white paper on artificial intelligence. In: EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies-The AI Act, pp. 211–280. Springer (2023)
37. Nnadi, L.C., Watanobe, Y., Rahman, M.M., John-Otumu, A.M.: Prediction of students' adaptability using explainable ai in educational machine learning models. Applied Sciences **14**(12), 5141 (2024)

38. Olari, V., Romeike, R.: Data-related practices for creating artificial intelligence systems in k-12. In: Proceedings of the 19th WiPSCE Conference on Primary and Secondary Computing Education Research. pp. 1–10 (2024)
39. Philipp Schmidt, F.B., Teubner, T.: Transparency and trust in artificial intelligence systems. Journal of Decision Systems **29**(4), 260–278 (2020). https://doi.org/10.1080/12460125.2020.1819094, https://doi.org/10.1080/12460125.2020.1819094
40. Rohlfing, K.J., Cimiano, P., Scharlau, I., Matzner, T., Buhl, H.M., Buschmeier, H., Esposito, E., Grimminger, A., Hammer, B., Häb-Umbach, R., et al.: Explanation as a social practice: Toward a conceptual framework for the social design of ai systems. IEEE Transactions on Cognitive and Developmental Systems **13**(3), 717–728 (2020)
41. S, D.V.P.: How can we manage biases in artificial intelligence systems – a systematic literature review. International Journal of Information Management Data Insights **3**(1), 100165 (2023). https://doi.org/https://doi.org/10.1016/j.jjimei.2023.100165, https://www.sciencedirect.com/science/article/pii/S2667096823000125
42. Sahidon, M.A.D.: Bridging the automated machine learning transparency and usability gap for non-experts: a user-centered design and evaluation study. Tech. rep., University of Nottingham
43. Shneiderman, B.: Human-centered AI. Oxford University Press (2022)
44. Slade, S., Prinsloo, P.: Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist **57**(10), 1510–1529 (2013)
45. Solopova, V., Rostom, E., Cremer, F., Gruszczynski, A., Witte, S., Zhang, C., López, F.R., Plößl, L., Hofmann, F., Romeike, R., et al.: Papagai: Automated feedback for reflective essays. In: German Conference on Artificial Intelligence (Künstliche Intelligenz). pp. 198–206. Springer (2023)
46. Tzafilkou, K., Perifanou, M., Economides, A.A.: Negative emotions, cognitive load, acceptance, and self-perceived learning outcome in emergency remote education during covid-19. Education and information technologies **26**(6), 7497–7521 (2021)
47. Voigt, P., Von dem Bussche, A.: The eu general data protection regulation (gdpr). A Practical Guide, 1st Ed., Cham: Springer International Publishing **10**(3152676), 10–5555 (2017)