Ethical FRAPPE – an adapted draft framework for ethical AIED

Bhoomika Agarwal, Corrie Urlings, Giel van Lankveld and Roland Klemke

Open Universiteit, Valkenburgerweg 177, 6419 AT Heerlen, The Netherlands

Abstract

Artificial Intelligence (AI) is pervading our lives in numerous ways today. It is important to apply ethical principles to guide the development and usage of AI systems to prevent harms or discrimination through AI algorithms. This has led to various ethical regulations and guidelines being formed at the corporate, national and supra-national level. The EU AI Act classifies the usage of AI in education as 'high-risk' as "such systems may violate the right to education and training as well as the right not to be discriminated against and perpetuate historical patterns of discrimination" [1, p. 26]. However, there has been little attention paid to ethics in AI in Education (AIED) in literature and there is only one existing framework to ethically guide AIED. AIED ethics is complex as it has to combine both general AI ethics and the ethics of educational technology. We aim to create a theoretical framework for AIED, comprising implementation guidelines for developers and organizational users of AI in education. In this paper, an existing draft framework by Holmes et al. is adapted by using insights from literature in the ethics of AI, ethics of educational technology and ethics of AIED.

Keywords

Artificial Intelligence, Education, Ethics

1. Introduction

Artificial Intelligence (AI), once a buzzword, is now a reality. It is being used in many aspects of our lives including healthcare, transport, communication, agriculture, finance and education. The usage of AI in classrooms and in education is promising and provides opportunities to improve the education process with technological innovations. AI has been applied in educational contexts in automation of administrative processes and tasks, curriculum and content development, instruction, and students' learning processes [3]. AI systems have enabled early detection and redress of learning shortcoming by analyzing student data - thereby providing a more customized learning experience for students [3]. Over the past decade, the use of AI tools to support or enhance learning has grown exponentially [4]. In a recent literature review, Chen et al. looked at 20 years of AI in Education (AIED) from 2000 to 2019 and shared several relevant findings: (a) the domain of AIED has received increased interest in the last few years, owing to the positive effect of AI on learning performance; (b) there is an increase in AIED literature over the years, showing an increased scientific output; (c) AIED research is especially found in interdisciplinary journals with a dual focus on education and technology [5].

Ethics plays an important role in guiding the usage of AI in our lives. As defined by Potter Stewart, "Ethics is knowing the difference between what you have a right to do and what is right to do" [6]. It is important to ethically guide the development and usage of AI for several reasons. The primary reason is that AI is being increasingly integrated into our lives and therefore has the potential for widespread influence and direct control over people's lives. This means that it could negatively or unfairly impact numerous lives with far-reaching consequences. AI technologies are being developed at a high speed to automate tasks that are traditionally done by humans. The parties implementing the automation of tasks are at risk of not fully considering the ethical consequences in an effort to improve efficiency and save costs [7]. When such automated tasks involve any sort of decision-making by AI, the decisions can impact the personal well-being of individuals and have a potential for dangerous consequences.

The EU AI Act [1] classifies the usage of AI in education as 'high-risk' as "such systems may violate the right to education and training as well as the right not to be discriminated against and perpetuate historical patterns of discrimination" [1, p. 26]. In addition, ethics for AIED have not been discussed at the forefront of national AI policy strategies [8]. Schiff examined 24 national AI policy strategies from G-7 and OECD countries and other important global actors such as India, China, Russia, Singapore and Malta [8]. The author found that remarkable attention has been paid to AI ethics in general, but this did not imply that attention has been paid to ethics in AIED in particular. Schiff also noted that the missing role

Proceedings of the Doctoral Consortium of the Seventeenth European Conference on Technology Enhanced Learning, September 12–16, 2022, Toulouse, France

bhoomika.agarwal@ou.nl (B. Agarwal); corrie.urlings@ou.nl (C. Urlings); giel.vanlankveld@ou.nl (G. v. Lankveld); roland.klemke@ou.nl (R. Klemke)

0000-0002-7347-8465 (B. Agarwal)

of education as a sector is an anomaly because many of these national AI policy documents "discuss the use of AI not only for healthcare, but also for transportation, agriculture, finance, and many other sectors" [8]. In addition, of the 4-5 countries that discussed AIED as a tool for teaching, learning and educational administration, none of them commented on or discussed AIED ethics. This is a cause of concern as there is no consideration about the ethical approach to AIED among policymakers [8]. Until now, there exists only one framework for ethical AIED developed by 'The Institute for Ethical AI in Education' aimed at those making procurement and application decisions regarding AIED [9].

AIED ethics is complicated as it has to consider both general AI ethics and the ethics of educational technology. On the one hand, there is an overlap between the ethics of AI, ethics of educational technology and ethics of AIED - suggesting that they should draw inspiration from each other [2]. On the other hand, the usage of AIED systems raises concerns such as the autonomy of teachers, responsibility and accountability for decisions made by AIED systems, impact of potential discrimination by AIED systems through historical biases, explainability of AIED systems, etc [2]. Owing to these concerns raised by AIED systems, AIED ethics deserves attention and there is a need to develop an ethical framework for guiding AIED ethics that is targeted at developers and organizational users of AIED.

Keeping the limited attention to AIED ethics in mind, we aim to create an ethical framework for AIED using the Ethical FRAPPE - a set of high-level ethical principles for AIED that are derived in this paper. This paper aims to answer the following research question: "Can Ethical FRAPPE be used to construct an exhaustive ethical framework for AIED?" Multiple steps are necessary to answer this research question: (1) Define the properties and aspects of an exhaustive ethical framework from literature; (2) Identify the ethical principles that can be used to form an exhaustive ethical framework for AIED; (3) Identify current and possible future use-case scenarios that an ethical framework for AIED can be applied to, such that the framework can be future-proof and evolve as AI evolves. However, several of these steps are out of scope for this paper. In this paper, we focus on the second step. As part of the second step, we build upon an existing draft framework for ethical AIED by Holmes et al. using insights from literature. The other two steps are planned as part of the future work, as described in section 5.

2. Background

A framework for AIED should aim to combine both the ethics of AI and the ethics of educational technology into a single framework, considering the overlap between these two domains. Thereby, 2.1 looks into the ethics of AI and 2.2 looks into the ethics of educational technology individually. 2.3 examines the overlap between the above two domains and looks at existing AIED ethics frameworks.

2.1. Ethics of AI

The ethics of AI in general have been studied extensively and numerous frameworks and policies have been developed for AI ethics. The inventory of AI Ethics guidelines by the Algorithm Watch [10] comprises 167 different guidelines on a corporate, national and supra-national level. Among these, some frameworks are notable. The Asilomar AI principles developed by the Future of Life Institute [11] has been adopted by 1797 AI and Robotics researchers and 3923 others. Furthermore, the 'Ethics Guidelines for trustworthy AI' have been proposed by the European Union [12]. The guidelines are encompassed in the 'AI Act', which is a proposed European law to regulate the usage of AI [1].

Floridi et al. encouraged an ethical approach to AI to incorporate the benefits of AI and mitigate the potential harms caused by AI in a balanced way. The authors proposed AI4People – a framework formed by the synthesis of existing sets of principles produced by various reputable, multi-stakeholder organisations and initiatives [13]. Their framework comprised of five principles-beneficence, non-maleficence, autonomy, justice and explicability [13]. These 5 principles have a major overlap with the principles found by Jobin et al. in their scoping review of AI ethics guidelines comprising 84 documents [14].

While there is a growing body of AI ethics guidelines and frameworks that can be found in literature [14, 10], these initiatives have primarily produced high-level ethical principles, tenets, values and abstract requirements for AI development and deployment [15]. This principle-based approach towards AI is criticised due to its inability to deal with the complexity of issues raised by AI [15, 16]. More specifically, the high-level ethical principles do not translate into practice automatically with the tools presently available to developers [17]. With the high number of abstract guidelines proposed, 'ethics washing' is on the rise by technology companies [18]. 'Ethics washing' occurs when technical companies define ethical policies to maintain outward appearances without following the principles in practice [18]. A second reason for criticism stems from the principle-based approached being aimed at a range of stakeholders and are thereby often difficult to understand for specific groups of users [16].

Although the principle-based approach is criticized to be ineffective due to issues such as ethics washing,

it forms a good first step towards defining an ethical framework. Thereby, we begin by defining the high-level ethical principles in this paper. As part of future work, we adopt a similar approach as [17], in which we plan to define requirements from ethical principles for AIED and map them to design-based research (DBR) process instead, as elaborated in section 5. Armstrong et al. define DBR in an educational setting as "a research approach that engages in iterative designs to develop knowledge that improves educational practices" [19]. As DBR brings educational research closer to everyday practice, this methodology is increasingly being used in designing educational research [20].

2.2. Ethics of educational technology

As AIED ethics needs to consider the ethics of educational technology, ethical policies for educational technology are reviewed here.

Pardo and Siemens identified four principles to categorize the issues derived from privacy in educational data: transparency, student control over the data, security, and accountability and assessment [21]. As Learning Analytics (LA) is a sub-field of AIED that uses educational data to optimize learning, the ethics of AIED should consider the ethics of LA. LA is defined in the proceedings of the 1st International Conference on Learning Analytics and Knowledge as "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" [22]. Sclater developed a code of practice for LA that advises educational institutions on how to use LA ethically. The authors considered eight focus areas - ownership and control, consent, transparency, privacy, validity, access, action, adverse impact, stewardship [23]. In a recent literature review on the ethics of LA in higher education, Pargman and McGrath found that the top three ethical areas most in LA articles are transparency, privacy, and informed consent [24]. In the context of Dutch higher education, Engelfriet et al. developed a guide to LA that focuses on the protection of personal student data. Drachsler and Greller developed an eight point checklist named DELICATE that can serve as a reflection aid for ethical and privacy-supported LA. The DELICATE checklist comprises 8 checkpoints- "Determination, Explain, Legitimate, Involve, Consent, Anonymise, Technical and External" as a quality checklist to make stakeholders aware and guide them through the process.

The ethics of educational technology contains issues that are relevant to the domain of education. Issues relating to student autonomy and control over their data can have long-term effects on the future of students. There needs to be regulations regarding informed consent and privacy of students, interpretation and management of student data. There is a clear overlap between ethics of educational technology and ethics of AIED - suggesting that ethics of AIED should draw inspiration from the ethics of educational technology and should build on top of frameworks for ethics of educational technology.

2.3. Ethics of AIED

This section looks at existing frameworks and guidelines for AIED ethics.

The conversation revolving around ethics for AIED was started over 20 years ago by Aiken and Epstein with an aim to raise awareness of researchers while designing educational systems [27]. The authors set down 10 principles for AIED systems based on "The Golden Rule for Computers in Education: Teach others as you would like to be taught" [27].

The first ethical framework for AIED was developed by The Institute for Ethical AI in Education that involves designers and developers for AIED and sets down guidelines for them [9]. However, this framework is aimed at the decision makers during the process of procurement and the application of AIED. This framework focuses on defining high-level ethical principles without any implementation guidelines that are relatable to developers during the design of AIED systems. It contains the downsides of the principle-based approach to AI ethics in the form of a lack of translation into practice for developers.

Holmes et al. conducted a survey with 17 domain experts comprising 10 open questions to gauge expert opinion about ethics of AIED [2]. They examined the various aspects of ethics of AIED and concluded that "the ethics of AIED cannot be reduced to questions about data or computational approaches alone" [2] and needs to account for the ethics of education – including, but not limited to – the purpose of learning, choice of pedagogy, role of technology with respect to teachers and access to education [2]. The authors created a 'strawman draft' framework, shown in Figure 1, that identified three areas of focus: "the ethics of data, computational approaches and education" and emphasized the overlaps between these foci. The authors identified 3 levels of overlap in their 'strawman draft' framework . The first level comprised of three foci: "the ethics of data, computational approaches and education" while the second level comprised of the overlap between each pair of foci. These 2 layers form the 'known unknowns' while the overlap between these 3 foci formed the 'unknown unknowns' [2].

3. Methods

This paper aims to answer the following research question: "Can Ethical FRAPPE be used to construct an ex-

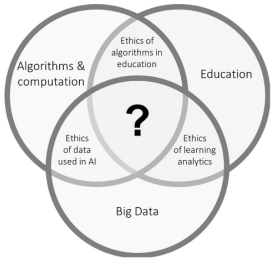

Figure 1: The 'strawman' draft framework for the ethics of AIED developed by [2]

haustive ethical framework for AIED?"

In order to answer this research question, the draft framework by Holmes et al. was selected as a foundational framework. This is because this 'strawman draft' framework is well-informed by experts in the domain of AIED and considers a template model for the essential aspects of ethical AIED. However, it only forms a skeleton model and does not contain the ethical principles involved in these domains. After making a few modifications, we fill in this gap in the 'strawman draft' framework by Holmes et al. by examining existing literature in the domains of both AI ethics, ethics of educational technology and ethics of AIED. High-level ethical principles are identified from literature and incorporated into this framework.

We proposed two modifications to the 'strawman draft' framework by Holmes et al.. Firstly, we elaborated on and defined the aspects in the intersection of these foci with an aim to throw light upon the 'known unknowns' and the 'unknown unknowns' stated by Holmes et al.. The 'strawman draft' framework defines the domains involved in ethics of AIED but does not elaborate on the ethical aspects of these domains. Thereby, we identified the ethical aspects involved in each of these foci based on literature, as shown in Figure 2.

Secondly, a huge overlap was noticed in the ethical aspects mapped to the foci of 'ethics of data' and 'ethics of computational approaches', as can be seen in Figure 2. Data and computational approaches were seen to be tightly coupled as any changes in one of them leads to changes in the other. For example, bias in data can lead to bias in the computational algorithm. Similarly, the interpretation and management of the data can have a direct effect on the privacy of the computational approach in the form of exposing sensitive attributes. Due to this tight coupling between the ethics of data and the ethics of computational approaches, they cannot be separated into 2 separate foci. Hence, we decided to combine them into a single focus. The revised and adapted version of our framework draft can be seen in Figure 3. It contains 2 focal areas: ethics of AI algorithms and ethics of educational technology, each containing corresponding ethical principles. The intersection of these 2 foci contains the ethical principles that form our theoretical framework.

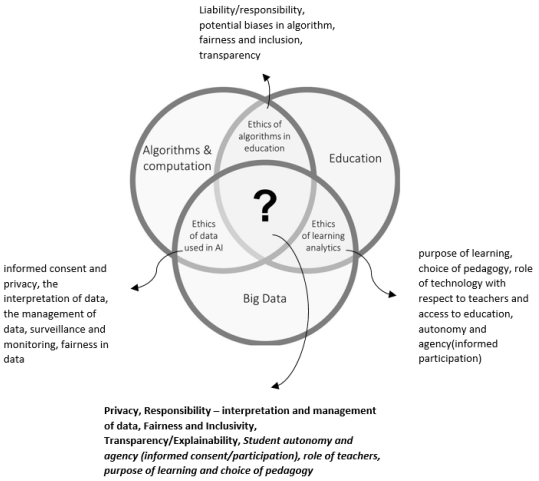

Figure 2: Revised version for the 'strawman' draft framework for the ethics of AIED adopted from [2]

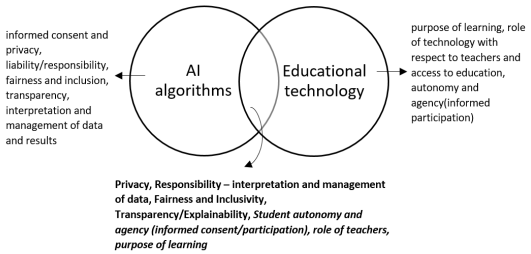

Figure 3: Adapted version for the 'strawman' draft framework for the ethics of AIED adopted from [2]

Following these modifications, literature in the domains of AI ethics, ethics of educational technology and ethics of AIED were reviewed. This was then used to obtain the ethical principles relevant to an ethical framework for AIED, abbreviated as the Ethical FRAPPE. The list of articles reviewed is grouped using the adapted draft framework as shown in Figure 3 into the ethics of AI, ethics of educational technology and ethics of AIED. Table 1 contains the list of selected articles that were reviewed to form the Ethical FRAPPE in order of year of publication. Following the adapted draft framework, these articles are grouped into the domains of ethics of AI, ethics of educational technology (EdTech) and ethics of AIED.

The high-level ethical principles seen in the literature

Table 1
List of selected articles

Year	Author(s)	Domain
2000	Aiken and Epstein [27]	Ethics of AIED
2014	Pardo and Siemens [21]	Ethics of EdTech
2016	Drachsler and Greller [26]	Ethics of EdTech
2016	Sclater [23]	Ethics of EdTech
2017	Engelfriet et al. [25]	Ethics of EdTech
2017	Prinsloo and Slade [29]	Ethics of EdTech
2017	Boddington [30]	Ethics of AI
2018	Floridi et al. [13]	Ethics of AI
2018	Whittaker et al. [7]	Ethics of AI
2019	Mittelstadt [15]	Ethics of AI
2019	Dignum [31]	Ethics of AI
2019	Jobin et al. [14]	Ethics of AI
2019	Crawford et al. [32]	Ethics of AI
2019	Kitto and Knight [33]	Ethics of EdTech
2019	Commission et al. [12]	Ethics of AI
2020	Morley et al. [17]	Ethics of AI
2020	Hagendorff [34]	Ethics of AI
2020	AlgorithmWatch [10]	Ethics of AI
2020	Vincent-Lancrin and van der Vlies [35]	Ethics of AIED
2021	Ryan and Stahl [16]	Ethics of AI
2021	Li et al. [36]	Ethics of AI
2021	Commission et al. [1]	Ethics of AI
2021	Miao et al. [4]	Ethics of AIED
2021	The Institute for Ethical AI in Education [9]	Ethics of AIED
2021	Holmes et al. [2]	Ethics of AIED
2021	Schiff [8]	Ethics of AIED
2021	Baker and Hawn [37]	Ethics of EdTech
2021	Pargman and McGrath [24]	Ethics of EdTech

were compared to each other. The ethical principles seen in a majority of the articles in each domain are identified and consolidated to create the Ethical FRAPPE. The 6 ethical principles identified as part of the Ethical FRAPPE are:

1. Fairness
2. Responsibility
3. Autonomy
4. Privacy
5. Purpose of learning
6. Explainability

Despite the presence of a large body of ethical guidelines, these guidelines rely on context-specific keywords and there exist multiple definitions of the ethical principles and technical terms involved [28]. This makes it challenging to interpret and operationalize these ethical values [28]. Keeping in mind the need for a common vocabulary to avoid misinterpretation of the ethical principles [28], we define the ethical aspects and explain them in the light of AIED ethics as below.

3.1. Fairness

Fairness, or Freedom from Bias is defined as - "Systematic unfairness perpetrated on individuals or groups, including pre-existing social bias, technical bias, and emergent social bias" [38]. AIED systems should not be designed such that the algorithms develop historically unfair prejudices by ensuring fair data that is inclusive, representative of the target population and without inaccuracies [16]. Any conscious or unconscious biases that are incorporated into AI algorithms through the data analysis can have a negative impact on the rights of individual students [4]. AIED should strive towards equitable access to AI technologies for all, keeping in line with SDG 4 set down by the UNESCO - "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" [4].

3.2. Responsibility

"Responsible AI is concerned with the fact that decisions and actions taken by intelligent autonomous systems have consequences that can be seen as being of an ethical nature" [31]. In [31], the author states that Responsible

AI should follow 3 ethical principles:

1. Accountability: refers to the ability of the AI system to explain and justify its decisions
2. Responsibility: refers to the role of people with regards to the AI system
3. Transparency: refers to the capability of AI systems to "describe, inspect and reproduce the mechanisms through which AI systems make decisions" [31]

In the light of AIED, responsibility is required to ensure accountability of decisions, responsibility of the developers and maintainers of AI towards its users and transparency of data and purpose of the system.

3.3. Autonomy

Human autonomy is defined as "Refers to people's ability to decide, plan, and act in ways that they believe will help them to achieve their goals" [38]. Autonomy, also called 'agency' in some ethical guidelines, ensures that the users of the systems are informed actors and have full control over their own decisions when they interact with the AIED system [16]. In AIED systems, student autonomy is important to ensure that students understand the purpose of the system and have complete control over their personal data, including the right to opt out of such systems without negative consequences. Students should be informed about the data being collected about them and should be involved in any decisions made using such data. Teacher autonomy is equally important to ensure that the role of teachers is highlighted in the form of the human-in-the-loop in the AIED system. By allowing teachers to review and act upon the decisions made by autonomous AIED systems, teacher autonomy can be ensured and unfair decisions by the AIED system can be reduced. It is also important that the data collected about the teachers should not have an adverse effect on their role in the classroom. This can be ensured by ensuring both teacher and student autonomy in AIED systems.

3.4. Privacy

While there are multiple definitions of the term Privacy, we choose the consolidated definition from [38] - "a claim, an entitlement, or a right of an individual to determine what information about himself or herself can be communicated to others" [38]. In the context of AIED systems, privacy pertains to the sharing of private and confidential data with others. A huge amount of digital data is stored about students through their online activity and there is a need to regulate the access and ownership of this data so that it is only accessible to the concerned parties and is not repurposed for other uses. There is a fear of educational institutions and employers using 'old' data and the usage of student data for commercial purposes [35].

3.5. Purpose of learning

At the moment, AIED systems are being used as an application use-case of AI instead of being motivated by learning goals. AIED is partly taking the form of data scientists looking for a context where predictive modelling and other AI techniques can be applied [39]. There is a need to criticially examine the purpose of learning and the performance measures that this purpose of learning is being reduced to. It is important to keep in mind that "theories of learning cannot, after all, be 'discovered' by algorithms" [39]

3.6. Explainability

Explainibility is "understanding how an AI model makes its decision" [40]. AIED systems should be actively monitored to ensure accurate and reproducible results that can be explained with the data and algorithmic functionality [16]. AIED models should be built to be explainable by design (using partially or fully explainable models) or post-hoc explainability methods should be used in the case of black-box models that are not inherently explainable [40]. In the case of AIED models, it is necessary to ensure that the decisions taken by the algorithm are explainable to humans in order to avoid negative harms to students.

4. Conclusion

This paper aimed to answer the research question: "Can Ethical FRAPPE be used to construct an exhaustive ethical framework for AIED?" In order to answer this question, this paper aims to identify the high-level ethical principles that can be used to construct an exhaustive ethical framework for AIED. The existing 'strawman draft' framework for ethical AIED by Holmes et al. was adapted by adding high-level ethical principles that were identified from existing literature in the domains of ethics of AI, ethics of educational technology and ethics of AIED. The six high-level ethical principles identified and consolidated from literature are abbreviated in the form of the Ethical FRAPPE for AIED: Fairness, Responsibility, Autonomy, Privacy, Purpose of learning and Explainability. The 6 ethical principles in the Ethical FRAPPE were defined in the context of AIED systems to form the first outline of our theoretical framework for AIED.

5. Future Work

The construction of our framework and implementation guidelines will be conducted in 3 phases: 'theoretical framework', 'evaluation framework' and 'instantiation'.

In the first phase, a theoretical ethical framework will be developed for AIED. In order to define an exhaustive ethical framework for AIED, it is first essential to look at what comprises a good ethical framework. To answer this, a literature review will be conducted. This paper describes the first part of the first phase where an existing draft model for AIED ethics was adapted by identifying high-level ethical principles from literature. In the future work, these high-level ethical principles will be converted to requirements and then be used to create the theoretical framework in the form of a checklist that contains practical guidelines for developers of AIED. The expected theoretical framework will be a checklist comprising definitions, requirements, formula and guidelines for ethical principles. This first draft of the framework will be evaluated by experts in the domain of AIED for face validity and content validity.

In the second phase, a methodology will be developed to quantify the ethics of AIED applications based on the theoretical framework. We refer to this methodology as the 'evaluation framework'. This evaluation framework will provide quantification tools for the ethical principles integrated in the form of a pipeline that can check existing AI systems for ethical soundness and provide recommendations for improvement. First, a subset of the ethical principles from the theoretical framework will be identified as 'focus' principles based on their prominence and relevance. Following this, various tools will be examined to identify suitable quantification tools for the focus ethical principles. Lastly, there will be an evaluation of different technologies for the architecture, followed by design and implementation of the evaluation pipeline. This evaluation pipeline will receive the trained AI algorithm, input data and output data as inputs and will give an ethical score as an output. This ethical score will be calculated as the sum of individual scores for each ethical principle. The individual score for each ethical principle will be based on the implementation of the guidelines from the theoretical framework and will also contain recommendations for improvements. If the ethical score for a majority of the ethical principles (exact threshold to be decided based on the number of ethical principles) is above 80%, the ethical evaluation will be passed. Such an ethical score allows for some trade-offs between principles in the event of conflicts between them, while ensuring that the system is ethical as a whole.

In the third phase, called 'instantiation', a proof of concept or instantiation of the evaluation framework will be developed. For this purpose, an AIED application will be developed which enables the identification of struggling students in a university, online, distance education setting. The main goal of this sample use case would be to improve teaching and learning processes on the whole and support teachers. The theoretical framework will be used to guide the design of this use case and the evaluation framework will be integrated into this AIED application to evaluate the ethics of this application. Additionally, the evaluation framework will be applied to some selected AIED models for evaluation such that they can cover various use cases. Finally, recommendations and guidelines will be provided for application of the theoretical and evaluation frameworks into other AIED applications. These recommendations will be developed for common challenges (such as biases or issues) seen in different classes/applications of reviewed AI algorithms from the literature. The applications of AI seen from literature will be grouped based on parameters such as the class of algorithms, coding language used and data type used.

References

[1] E. Commission, C. Directorate-General for Communications Networks, Technology, Proposal for a regulation of the european parliament and the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts (2021). URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206.

[2] W. Holmes, K. Porayska-Pomsta, K. Holstein, E. Sutherland, T. Baker, S. B. Shum, O. C. Santos, M. T. Rodrigo, M. Cukurova, I. I. Bittencourt, K. R. Koedinger, Ethics of AI in Education: Towards a Community-Wide Framework, International Journal of Artificial Intelligence in Education (2021). doi:10.1007/S40593-021-00239-1.

[3] L. Chen, P. Chen, Z. Lin, Artificial intelligence in education: A review, IEEE Access 8 (2020) 75264–75278. doi:10.1109/ACCESS.2020.2988510.

[4] F. Miao, W. Holmes, R. Huang, H. Zhang, et al., AI and education: A guidance for policymakers, UNESCO Publishing, 2021.

[5] X. Chen, D. Zou, H. Xie, G. Cheng, C. Liu, Two decades of artificial intelligence in education, Educational Technology & Society 25 (2022) 28–47.

[6] P. Stewart, Ethics and problem solving 11 (2021).

[7] M. Whittaker, K. Crawford, R. Dobbe, G. Fried, E. Kaziunas, V. Mathur, S. M. West, R. Richardson, J. Schultz, O. Schwartz, AI now report 2018, AI Now Institute at New York University New York, 2018.

[8] D. Schiff, Education for AI, not AI for Education: The Role of Education and Ethics in National AI Policy Strategies, International Journal of Artificial

Intelligence in Education (2021) 1–37. doi:10.1007/S40593-021-00270-2/.

[9] The Institute for Ethical AI in Education, The Ethical Framework for AI in Education, Technical Report, The Institute for Ethical AI in Education, 2021. URL: https://www.buckingham.ac.uk/research-the-institute-for-ethical-ai-in-education/.

[10] AlgorithmWatch, AI Ethics Guidelines Global Inventory - AlgorithmWatch, 2020. URL: https://algorithmwatch.org/en/ai-ethics-guidelines-global-inventory/.

[11] Future of Life Institute, Asilomar AI Principles, 2017. URL: https://futureoflife.org/2017/08/11/ai-principles/.

[12] E. Commission, C. Directorate-General for Communications Networks, Technology, Ethics guidelines for trustworthy AI, Publications Office, 2019. doi:doi/10.2759/346720.

[13] L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Schafer, P. Valcke, E. Vayena, AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations, Minds and Machines 28 (2018) 689–707. doi:10.1007/S11023-018-9482-5/.

[14] A. Jobin, M. Ienca, E. Vayena, Artificial Intelligence: the global landscape of ethics guidelines, Nature Machine Intelligence 1 (2019) 389–399. doi:10.1038/s42256-019-0088-2.

[15] B. Mittelstadt, Ai ethics–too principled to fail, arXiv preprint arXiv:1906.06668 (2019).

[16] M. Ryan, B. C. Stahl, Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications, Journal of Information, Communication and Ethics in Society 19 (2021) 1477–996. URL: http://creativecommons.org/licences/by/4.0/legalcode. doi:10.1108/JICES-12-2019-0138.

[17] J. Morley, L. Floridi, L. Kinsey, A. Elhalal, From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices, Science and Engineering Ethics 26 (2020) 2141–2168. doi:10.1007/S11948-019-00165-5/.

[18] E. Bietti, From ethics washing to ethics bashing: A view on tech ethics from within moral philosophy, in: Proceedings of the 2020 conference on fairness, accountability, and transparency, 2020, pp. 210–219.

[19] M. Armstrong, C. Dopp, J. Welsh, Design-Based Research - The Students' Guide to Learning Design and Research, EdTech Books, 2018, pp. 40–45. URL: https://edtechbooks.org/studentguide/design-based_research.

[20] D.-B. R. Collective, Design-based research: An emerging paradigm for educational inquiry, Educational researcher 32 (2003) 5–8.

[21] A. Pardo, G. Siemens, Ethical and privacy principles for learning analytics, British Journal of Educational Technology 45 (2014) 438–450. doi:10.1111/BJET.12152.

[22] P. Long, G. Siemens, G. Conole, D. Gašević, LAK '11: Proceedings of the 1st International Conference on Learning Analytics and Knowledge, Association for Computing Machinery, New York, NY, USA, 2011.

[23] N. Sclater, Developing a code of practice for learning analytics, Journal of Learning Analytics 3 (2016) 16–42. URL: https://learning-analytics.info/index.php/JLA/article/view/4512. doi:10.18608/jla.2016.31.3.

[24] T. C. Pargman, C. McGrath, Mapping the terrain of ethics in learning analytics: A systematic literature review of empirical research, Journal of Learning Analytics (2021).

[25] A. Engelfriet, J. Manderveld, E. Jeunink, Learning analytics under the (dutch) personal data protection act. policy paper. surfnet, 2017. URL: https://www.surf.nl/en/guide-to-learning-analytics-under-the-personal-data-protection-act.

[26] H. Drachsler, W. Greller, Privacy and Analytics-it's a DELICATE Issue A Checklist for Trusted Learning Analytics, Proceedings of the Sixth International Conference on Learning Analytics & Knowledge (2016) 89–98. doi:10.1145/2883851.

[27] R. M. Aiken, R. G. Epstein, Ethical Guidelines for AI in Education: Starting a Conversation, International Journal of Artificial Intelligence in Education 11 (2000) 163–176. URL: http://www.cs.wcupa.edu/~epstein/AIStories.html.

[28] A. Theodorou, V. Dignum, Towards ethical and socio-legal governance in ai, Nature Machine Intelligence 2 (2020) 10–12.

[29] P. Prinsloo, S. Slade, Ethics and Learning Analytics: Charting the (Un)Charted, Handbook of Learning Analytics (2017) 49–57. URL: http://oro.open.ac.uk/49456/. doi:10.18608/HLA17.004.

[30] P. Boddington, Towards a code of ethics for artificial intelligence, Springer, 2017.

[31] V. Dignum, Responsible artificial intelligence: how to develop and use AI in a responsible way, Springer Nature, 2019.

[32] K. Crawford, R. Dobbe, T. Dryer, G. Fried, B. Green, E. Kaziunas, A. Kak, V. Mathur, E. McElroy, A. N. Sánchez, et al., Ai now 2019 report, New York, NY: AI Now Institute (2019).

[33] K. Kitto, S. Knight, Practical ethics for building learning analytics, British Journal of Educational Technology 50 (2019) 2855–2870.

[34] T. Hagendorff, The Ethics of AI Ethics: An Evaluation of Guidelines, Minds and Machines 30 (2020) 99–120. doi:10.1007/S11023-020-09517-8/.

[35] S. Vincent-Lancrin, R. van der Vlies, Trustworthy artificial intelligence (ai) in education: Promises and challenges (2020).

[36] B. Li, P. Qi, B. Liu, S. Di, J. Liu, J. Pei, J. Yi, B. Zhou, Trustworthy ai: From principles to practices, arXiv preprint arXiv:2110.01167 (2021).

[37] R. S. Baker, A. Hawn, Algorithmic bias in education, International Journal of Artificial Intelligence in Education (2021) 1–41.

[38] B. Friedman, P. H. Kahn, A. Borning, A. Huldtgren, Value Sensitive Design and Information Systems, Springer Netherlands, Dordrecht, 2013, pp. 55–95. URL: https://doi.org/10.1007/978-94-007-7844-3_4. doi:10.1007/978-94-007-7844-3_4.

[39] C. Perrotta, N. Selwyn, Deep learning goes to school: toward a relational understanding of ai in education, Learning, Media and Technology 45 (2020) 251–269. doi:10.1080/17439884.2020.1686017.

[40] B. Li, P. Qi, B. Liu, S. Di, J. Liu, J. Pei, J. Yi, B. Zhou, Trustworthy ai: From principles to practices, arXiv preprint arXiv:2110.01167 (2021). URL: https://doi.org/10.48550/arXiv.2110.01167.