No: 02 2026
https://doi.org/10.64400/ea26-02-07
ISSN 2224-5529 2224-0829 2223-5124 2222-940X

NAXÇIVAN DÖVLƏT UNİVERSİTETİNİN ELMİ ƏSƏRLƏR JURNALI

İCTİMAİ ELMLƏR

EXPLORING FRAMEWORKS FOR THE ETHICAL DEPLOYMENT OF ARTIFICIAL INTELLIGENCE IN HIGHER EDUCATION

Anyinyo Norman OYUGA
Baku State University, Baku, Azerbaijan
**Yazışılan müəllif:* anyinyonorman@gmail.com

NƏŞR TARİXİ:
Qəbul edilmə tarixi: 06.06.2026
Nəşr edilmə tarixi: 25.06.2026

KEYWORDS:
Artificial Intelligence, Higher Education, Framework, Ethics, Algorithmic Bias.

SUMMARY
In the 21st century, Artificial Intelligence (AI) has emerged as one of the most transformative technologies in various sectors, including higher education, where it is revolutionising core functions. Existing research shows persistent gaps in how universities govern and ethically integrate AI, citing inconsistent or opaque governance structures that cause uncertainty to staff and students. To fill this gap, this study explored frameworks that higher education institutions can deploy to ensure that ethical standards are upheld when using artificial intelligence. A qualitative literature review approach that involved examination of secondary data published in journal articles was administered. Findings from the literature search revealed that the main possible frameworks that can govern artificial intelligence use in higher learning establishments include decision-making, participatory, data security, privacy, accountability, and equity-centric frameworks. The results will be a useful resource for policymakers, researchers, educators and institutional managers. This paper recommends further stakeholder-segmented research to evaluate the differential impacts of AI ethics policies and tools.

INTRODUCTION

The recent years have witnessed an astronomical expansion in artificial learning capabilities, especially machine learning and other deep learning algorithms. This advanced use of technology has led to adverse implications on industrial, social, environmental and economic changes on humans (Hu, 2022). Before, artificial intelligence (AI) was understood as a simulation of human intelligence, but more advanced descriptions have emerged lately. According to Dignum (2019), AI is "a socio-technical ecosystem, recognizing the interaction between people and technology, and how complex infrastructures affect and are affected by society and by human behaviour". This definition not only recognises AI as a technology for automation of processes and tasks but also acknowledges AI's impacts on societal and human interactions and its influence on the ecosystem.

As much as AI can expand the space for employment opportunities, there is still a lack of readiness among the current workforce to take up new roles, and this implies that even the future workforce will need to be equipped with special skills to work with AI (Forum, 2018). Research has recommended complementing and augmenting human workers by using AI, which has led to the advocacy of AI-worker partnership and collaboration (Karakose et al., 2023). Pressure continues to pile across all sectors of society, including in education, for the adequate preparation of current and future workforces to apply AI in non-labour-intensive tasks like art creation and decision-making (Kong et al.,2024). Preparation of a future workforce that is technologically savvy requires AI literacy, which is defined as "the elements that the workforce needs to harness AI and form a synergistic relationship with the technology" (Kong et al., 2024). Notwithstanding the benefits accompanying AI adoption, myriad issues have formed part of the contemporary issues under discussion: displacement of workers, exacerbated inequality, lack of human autonomy, cybercrimes, misinformation, disinformation, AI-driven biases, unfairness and inexplicability (Nguyen & Hekman, 2024).

Studies suggest that AI literacy should be promoted universally to all students, but most institutions have targeted computer science students, and this undermines the appreciation of real-life

problems in society (Kong et al.,2024). There is a consensus on the classification of AI literacy into 4 dimensions: cognitive, meta-cognitive, affective and social (Kong et al.,2024; Wang et al.,2023). The social dimension of this classification relates to the ethical considerations surrounding solutions based on AI. This dimension informs the students on the scope of using AI for problem-solving (Kong et al.,2024). Unlimited access to powerful generative AI, for instance ChatGPT, by the general public recently, has increasingly become a point of concern among researchers, corporate leaders and policy makers because of its ramifications on AI ethics and profound risks to human society (Duffy & Maruf, 2023). Because of the reality of the risks the world is facing, the use of AI needs to be applied ethically and responsibly. Although ethics and responsibility are commonly used interchangeably, Dignum (2019) points out a narrow distinction, stating that ethics is concerned with morals and values, while responsibility encompasses the practicality of applying ethical concerns and legal, economic and cultural ones for the benefit of all.

Due to AI's role in the digital space and its influence on personalities and society, it is imperative to have a responsible approach to AI so that AI technologies are not entirely blamed for decisions and actions where humans have a role in designing and deploying systems that are in tandem with human values and ethics and contribute to welfare and sustainability (Dignum, 2019). To promote responsible AI development, governments and non-governmental organisations, such as the U.S. Executive Order and the European Union AI Act, have formulated regulatory, policy, and educational initiatives to formalise responsible AI guidelines (European Union, 2024; The White House, 2023). Alternatively, both educational and non-educational settings in Finland have witnessed public AI literacy initiatives through massive open online courses (MOOCs); the Singaporean government is supporting its nationals with a dedicated AI ethics course in governance and education; world-class universities like Massachusetts Institute of Technology, Stanford University, and Purdue University have implemented programs promoting responsible AI use (AI Singapore, 2024; European Commission, 2021; Wiese et al., 2025).

Recently, research and academia in general have suffered from credibility issues ranging from plagiarism and unethical use of AI. Despite the innovation of anti-plagiarism software and techniques of detecting malpractice, there is a sense of inadequacy as far as the deterrence of governance structures is concerned. Based on heightened advocacy by governments to get their citizens AI-literate by systematically using policy and regulatory frameworks and world-class universities promoting responsible AI use, this study seeks to add to the body of knowledge on the useful regulatory frameworks that can be implemented to facilitate ethical AI use. The following question guided the study: What possible frameworks can higher education institutions utilize to ensure the ethical deployment of Artificial Intelligence?

MATERIAL AND METHODS

This study used a qualitative approach that involved a literature review from already published reviews and research articles related to the topic. The methodology offers the depth, interpretive flexibility, and conceptual sensitivity necessary for examining emerging and ethically complex topics, for instance, the deployment of AI in higher education. Qualitative reviews, unlike strictly systematic or meta-analytic methods, permit the researcher to explore the rapidly expanding body of literature and correlate fragmented themes, assumptions and generalizations. Additionally, qualitative syntheses support the development of new conceptual and theoretical models by critically interpreting novel literature (Younas et al., 2023). Hence, the researcher found it possible to analyse complex phenomena by organising his findings thematically within a short span of time. More recent studies that are less than 10 years old were preferred because the question under investigation is rapidly evolving. The findings from these secondary sources were carefully examined and rigorously analysed to inform the discussion. The main databases used to search for data were ScienceDirect and Google Scholar.

RESULTS

Modern society is increasingly using artificial intelligence (AI) systems in making decisions that have a direct implication on human lives; banking and commerce, criminal predictions and employment. Due to the possibility of discrimination and employment biases in case handling, institutions of higher learning need a framework for mitigating these biases. A decision-making framework for mitigating algorithmic biases ensures that these biases are identified and rectified to promote inclusivity and equity. To minimize discriminatory outcomes, diverse datasets should be used to train AI systems, and before deployment, multi-disciplinary experts need to test and validate the tools (Rubin, 2024). It is proven that potential biases of AI outcomes can be eliminated through quality data training and labelling (Xu et al., 2025).

Higher education institutions need to formulate policies stipulating what ethical usage of AI in the institution entails. A participatory framework for ethical usage should be developed based on a compilation of issues that have arisen previously, so that they are mitigated or eradicated. Nevertheless, these guidelines on ethical usage must pass the test of fairness, accountability, data privacy and transparency and undergo regular reviews to accommodate advancements in technology. AI should not be handled exclusively by tech companies; it should be done consultatively with higher education institutions and other stakeholders so that their input is considered. This has not always been the case since most developers compete to outsmart one another. Tech companies and universities should work within a participatory framework that can co-create ethical solutions. The framework should compel tech companies to invest in research on ways of ethically applying AI in higher education (Dabis & Csáki, 2024).

Privacy and security are primarily data governance procedures, which encompass collection, storage, access, usage and management, particularly when it involves students and educators (Fu & Weng, 2024). A data security and privacy policy framework would regulate the usage of AI for administrative, research, teaching and learning purposes. Being universally accessible, AI may threaten the safety of governmental or institutional information if not kept in check. A data security and privacy policy framework helps prevent unnecessary publicity and limits the chances of disseminating classified data to competitors and adversarial entities, as it is the responsibility of educational institutions and cloud service providers to guarantee the security of students' and educators' data (Fu & Weng, 2024). Due to a few legal uncertainties the users of AI face, policymakers need to ensure that integrity is upheld when setting standards that pertain to the appropriate use of AI in education and research (Mittelsteadt, 2023).

An accountability framework that allows for establishing governance boards to oversee implementation and transparency is beneficial. Accountability is defined as “the fact or state of taking responsibility for your decisions or actions so that you can explain them if necessary” (Oxford University Press, n.d., para.1). An openly accessible AI platform that cannot be audited is detrimental; thus, a reporting mechanism is useful in tracking and reprimanding abusers. As AI applications continue to report positive performance, corresponding inventions are being adopted by the industry and patent registrations are improving. Small use cases are morphing into bigger and more intricate systems that impact lives directly. Unlike earlier technological developments, substantial decisions can be made without significant human involvement, bringing to the fore a key concern about how such systems can be held accountable. The two facets of accountability that must be factored in when drafting a policy framework influenced by AI tools are explainability and liability (Boch et al., 2022). Furthermore, achieving explainability dictates that the custodian of the data should be answerable to the people holding the right to explanation. Contrastingly, scholars have debated whether the law gives a general right to explanation for AI processes (Bibal et al.,2021). Liability is “the state of being legally responsible for something” (Oxford University Press, n.d., “Definition 1”).

In AI education, fairness and equity are critical areas of focus since they are connected to education quality (Fu & Weng, 2024). On examining pertinent issues that plague higher education, inequity in accessing resources features prominently. Students from marginalized backgrounds encounter obstacles to funding, academic support and networking opportunities. In the long run, their academic performance and career prospects are reduced. AI can help develop an equity-centric framework critical for bridging the gap of the digital divide in higher education through analytics and the study of trends (Reimer & Hill, 2024). In addressing equity and fairness, specific student contexts should be traced rather than homogenising their

needs and ignoring individual differences (Fu & Weng, 2024). During the publication of the National Education Technology Plan by the United States Department of Education Office of Educational Technology (ED), the plan cited three digital divides: the digital use divide, the digital design divide and the digital access divide. The digital use divide describes how students who have been able to use technology with higher-order thinking skills like coding and critical thinking differ from those who have not. The digital design divide refers to the gap between the school systems that provide time for teachers to develop digital learning materials and those that do not. Finally, the digital access divide is the distinction between those to whom digital devices, the internet and instructional content are accessible and those to whom they are not (Reimer & Hill,2024).

DISCUSSION

The increasing reliance of modern society on AI has exacerbated the risk of algorithmic bias. Higher education institutions (HEIs), therefore, need a robust decision-making framework to identify, mitigate and deter such discrimination to foster inclusivity and fairness in AI-driven outcomes. There is further scientific proof that associates sophisticated data curation and careful labelling with significant reduction or elimination of plausible biases in AI systems (Xu et al., 2025). Secondly, HEIs need to develop clear institutional policies that stipulate the parameters for acceptable use of AI. It is the analysis of past ethical challenges aimed at mitigating and eliminating recurrent risks that informs the formulation of a participatory framework. Therefore, the role of enforcing ethical AI practises must not be surrendered to technology companies alone; other actors and stakeholders, including universities, who are significant consumers, must be onboarded. Dabis & Csáki (2024) posit that for a participatory framework to be considered as effective, it needs to support collaborative development of ethical solutions and enable tech companies to channel meaningful investments in researching AI applications in higher education.

Thirdly, the findings reveal that privacy and security in AI deployment are dependent on a robust data governance practice, including collation, storage, access, utilization and management of lecturers and students data (Fu & Weng, 2024). The rationale for having a data security and privacy policy framework is to ensure responsiveness in administrative, research, teaching and learning activities. Unrestrained deployment of AI would pose several risks to the confidentiality of institutional and governmental information since the use of AI has become increasingly prevalent. Such a framework mitigates such risks by averting unauthorised exposure of classified data to adversaries or competitors. Eventually, the responsibility for safeguarding students' and educators' data rests with the educational institution as well as the cloud service providers. Given the legal bottlenecks around AI use, policymakers must develop concise regulations to ensure that integrity and ethics are upheld in research and educational contexts with respect to data (Mittelsteadt, 2023).

This review emphasised a strong accountability framework for AI integration in higher education, including the creation of governance boards for oversight and transparency. With the shift from small-scale AI use to more complex decision-making tools, the aspect of accountability is vital. Unlike previous technological advancements, AI requires very limited human intervention for decision-making, hence the discourse about responsibility remains unsettled. Therefore, two core dimensions of accountability, which are explainability and liability, require addressing using an effective policy framework (Boch et al., 2022). The principle of explainability demands that custodians of data remain answerable to stakeholders, though the legality of this perspective is still under contest (Bibal et al., 2021). Furthermore, regular auditing and transparent reporting systems aid in tracking usage, deterring misuse and promoting trust in AI applications.

Fairness and equity are foundational dimensions of AI-enhanced education as they reflect the wider expectations of educational quality. In higher education, the problem of resource-access inequity is a persistent structural challenge since unevenly distributed opportunities shape institutional effectiveness and student learning outcomes. Disadvantaged students face barriers like securing funding, academic support, networking opportunities and most importantly, a long-term reduction in academic performance and career prospects among the digitally disadvantaged students who fail to benefit from AI-enabled educational tools. Recent studies show the ability of AI to

contribute to equity-centric frameworks that not only identify systematic gaps and generate data-driven insights but also suggest interventions that help mitigate inequalities in higher education. Furthermore, the results show that addressing equity and fairness concerns in educational settings needs to take into account learning contexts rather than relying on generalisations. This aligns with the view to recognise variances across learner demographics to ensure that evaluative practises conform to contexts. An equity framework ensures that the stratification caused by the digital access divide in higher education, which bars access to devices, connectivity and digital instructional content, is reduced or eliminated.

CONCLUSION

The emergence of AI impacted education in ways that had not been anticipated before. In response, the higher education sector is gradually formulating frameworks that are helpful for the ethical and proper use of this technology. Five key frameworks were identified as the most vital for ensuring ethical application of AI in higher education. A decision-making framework is useful in identifying, mitigating and deterring algorithmic biases to promote objectivity in decisions arrived at. Participatory framework ensures that all actors and stakeholders are involved in the designing and curation of content and devices used in institutions, and that all parties are liable in case of any eventuality. Data security and privacy framework governs the risks of exposure of individual data to unauthorised parties who may misuse it, as data is supposed to be confidential unless the handler has the legal mandate to expose or share it. Accountability framework guides how custodians of data are required to bear the responsibility of answering to stakeholders. Lastly, this study states that an equity-centric framework helps to identify systematic gaps and data-driven insights on interventions to support the digitally disadvantaged to mitigate inequalities in higher education. AI is rapidly evolving, and the number of frameworks is increasing exponentially with time. This study suggests further stakeholder-segmented research to evaluate the differential impacts of AI ethics policies and tools.

REFERENCES

AI Singapore. (2024). AI Singapore. https://aisingapore.org/.

Bibal, A., Lognoul, M., De Streel, A., & Frénay, B. (2021). Legal requirements on explainability in machine learning. Artificial Intelligence and Law, 29, 149-169.

Boch, A., Hohma, E., & Trauth, R. (2022). Towards an accountability framework for AI: Ethical and legal considerations. Institute for Ethics in AI, Technical University of Munich: Munich, Germany. https://ieai.mcts.tum.de/

Dabis, A., & Csáki, C. (2024). AI and ethics: Investigating the first policy responses of higher education institutions to the challenge of generative AI. Humanities and Social Sciences Communications, 11(1), 1-13. https://doi.org/10.1057/s41599-024-03526-z

Dignum, V. (2019). Responsible artificial intelligence: how to develop and use AI in a responsible way (Vol. 2156). Cham: Springer. https://doi.org/10.1007/978-3-030-30371-6

Duffy, C., & Maruf, R. (2023). Elon Musk warns AI could cause 'civilization destruction' even as he invests in it. CNN Business.

European Commission. (2021). Finland AI strategy report. https://ai-watch.ec.europa.eu/countries/finland/finland-ai-strategy-report_en

European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence. http://data.europa.eu/eli/reg/2024/1689/oj/eng.

Forum, W. E. (2018). The future of jobs report 2018. Insight Report, Centre for the New Economy and Society, Geneva/Switzerland.

Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. Computers and Education: Artificial Intelligence, 100306. https://doi.org/10.1016/j.caeai.2024.100306

https://doi.org/10.1080/0144929X.2022.2072768
Hu, C. (2022). How artificial intelligence exploded over the past decade. Popular Science.
Xu, Q., Wu, Y., Zheng, H., Yan, H., Wu, H., Qian, Y., ... & Liu, B. (2025). Standardization in artificial general intelligence model for education. Computer Standards & Interfaces, 94, 104006. https://doi.org/10.1016/j.csi.2025.104006
Karakose, T., Demirkol, M., Aslan, N., Köse, H., & Yirci, R. (2023). A conversation with ChatGPT about the impact of the COVID-19 pandemic on education: Comparative review based on human–AI collaboration. Educational Process: International Journal, 12(3), 7.
Kong, S. C., Cheung, M. Y. W., & Tsang, O. (2024). Developing an artificial intelligence literacy framework: Evaluation of a literacy course for senior secondary students using a project-based learning approach. Computers and Education: Artificial Intelligence, 6, 100214. https://doi.org/10.1016/j.caeai.2024.100214
Mittelsteadt, M. (2023). Artificial intelligence: An introduction for policymakers. Mercatus Research Paper. https://ssrn.com/abstract=4361563
Nguyen, D., & Hekman, E. (2024). The news framing of artificial intelligence: A critical exploration of how media discourses make sense of automation. AI & Society, 39(2), 437–451. https://rdcu.be/fgf95
Oxford University Press. (n.d). Liability. In Oxford English Dictionary. Retrieved April 8, 2026.
Oxford University Press. (n.d.). Accountability, n. In Oxford English Dictionary of Academic English. Retrieved April 8, 2026, from https://doi.org/10.1093/OED/2470482822
Reimer, T., & Hill, J. (2024). Preparing Post-Pandemic, Equity-Focused Educational Leaders: Technology Requires Administrators to Reimagine Schools. International Journal of Educational Leadership Preparation, 19(1), 138-156.
Rubin, L. M. (2024). ChatGPT and Research Ethics. ChatGPT and Global Higher Education: Using Artificial Intelligence in Teaching and Learning, 179.
The White House. (2023). Executive order on the safe, secure, and trustworthy development and use of artificial intelligence. The White House. https://www.whiteh ouse.gov/briefing-room/presidential-actions/2023/10/30/executive-orderon-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/
Wang, B., Rau, P.-L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. Behaviour & Information Technology, 42(9), 1324–1337.
Wiese, L. J., Patil, I., Schiff, D. S., & Magana, A. J. (2025). AI Ethics Education: A Systematic Literature Review. Computers and Education: Artificial Intelligence, 100405.
Younas, A., Fàbregues, S., Durante, A., Escalante, E. L., Inayat, S., & Ali, P. (2023). Proposing the "MIRACLE" narrative framework for providing thick description in qualitative research. International Journal of Qualitative Methods, 22, 16094069221147162. https://doi.org/10.1177/16094069221147162

XÜLASƏ

ALİ TƏHSİLDƏ SUNİ İNTELEKTİNİN ETİK TƏTBİQİ ÜÇÜN MÜMKÜN ÇƏRÇİVƏLƏR

Anyinyo Norman Oyuga.
Bakı Dövlət Universiteti, Bakı, Azərbaycan

21-ci əsrdə Süni İntellekt (AI) əsas funksiyaları dəyişdirən ali təhsil də daxil olmaqla müxtəlif sektorlarda ən transformativ texnologiyalardan biri kimi ortaya çıxdı. Mövcud tədqiqatlar universitetlərin süni intellekti idarə etməsi və etik şəkildə inteqrasiya etməsi sahəsində davamlı boşluqların olduğunu göstərir ki, bu da qeyri-ardıcıl və ya qeyri-şəffaf idarəetmə strukturlarına istinad

edərək işçilər və tələbələr arasında qeyri-müəyyənlik yaradır. Bu boşluğu doldurmaq üçün bu tədqiqat ali təhsil müəssisələrinin süni intellektdən istifadə edərkən etik standartların təmin edilməsini təmin etmək üçün tətbiq edə biləcəyi çərçivələri araşdırdı. Jurnal məqalələrində dərc edilmiş ikinci dərəcəli məlumatların araşdırılmasını əhatə edən keyfiyyətli tədqiqat yanaşması tətbiq edilmişdir. Ədəbiyyat axtarışından əldə edilən nəticələr göstərir ki, ali təhsil müəssisələrində süni intellektdən istifadəni idarə edə biləcək əsas mümkün çərçivələrə qərar qəbuletmə, iştirakçılıq, məlumat təhlükəsizliyi, məxfilik, hesabatlılıq və kapital mərkəzli çərçivələr daxildir. Nəticələr siyasətçilər, tədqiqatçılar, müəllimlər və institusional menecerlər üçün faydalı mənbə olacaq. Bu araşdırma, AI etikası siyasətlərinin və alətlərinin müxtəlif maraqlı tərəflər üzrə fərqli təsirlərini qiymətləndirmək məqsədilə maraqlı tərəflərə bölünmüş əlavə tədqiqatların aparılmasını tövsiyə edir.

Açar sözlər: *Süni intellekt, ali təhsil, çərçivə, etika, alqoritmik qərəz.*

АННОТАЦИЯ

ВОЗМОЖНЫЕ РАМКИ ЭТИЧЕСКОГО ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ВЫСШЕМ ОБРАЗОВАНИИ

Аниньо Норман Оюга
Бакинский государственный университет, Баку, Азербайджан

В 21 веке искусственный интеллект (ИИ) оказался среди самых преобразующих технологий в различных секторах, включая высшее образование, где он радикально изменяет основные функции. Существующие исследования показывают наличие устойчивых пробелов в том, как университеты управляют и этически интегрируют искусственный интеллект, ссылаясь на непоследовательные или непрозрачные структуры управления, которые вызывают неопределённость у сотрудников и студентов. Чтобы заполнить этот пробел, это исследование было направлено на изучение рамок, которые высшие учебные заведения могут применять для обеспечения соблюдения этических стандартов при использовании искусственного интеллекта. Качественный подход к исследованию, который включал изучение вторичных данных, опубликованных в журнальных статьях. Результаты поиска литературы показали, что основные возможные рамок, которые можно использовать для управления использованием искусственного интеллекта в высших учебных заведениях, включают рамку принятия решений, рамку участия, рамку безопасности и конфиденциальности данных, рамку подотчетности и рамку, ориентированную на равенство. Результаты станут полезным ресурсом для политиков, исследователей, педагогов и руководителей учреждений. Данная статья рекомендует проведение дополнительных исследований, дифференцированных по группам заинтересованных сторон, для оценки различного влияния политик и инструментов в области этики искусственного интеллекта

Ключевые слова: *Искусственный интеллект, высшее образование, рамочная основа, этика, Алгоритмическая предвзятость*