Crafting AI Terms of Use for Higher Education

Introduction

The use of AI for educational purposes has triggered a revolution in education (Chaudhry & Kazim, 2022; Yueh & Chiang, 2020). The increasing adoption of generative artificial intelligence (AI) in higher education underscores the necessity of formulating ethical guidelines that promote responsible and advantageous usage of these groundbreaking instruments by learners (Lim et al., 2022). Regrettably, the ethical dimensions of AI applications in education have been mostly disregarded in educational research (Yu & Yu, 2023) due to the rapid advancement in AI technology. Despite the challenges, educational researchers and higher education institutions have a responsibility to provide ongoing guidance so that AI technologies can be implemented in an ethical manner.

In pursuit of the stated objective, this paper begins by first encapsulating the salient principles outlined in IEEE's Ethically Aligned Design (EAD2v2) Standards (2018), which serves as a crucial reference for the ethical development and deployment of AI and autonomous systems. Then, the highly influential Fjeld et al.'s (2020) paper – a Berkman Klein Center's (a Harvard University research centre) publication is explained. Drawing from the insights offered by these two prominent frameworks, a comprehensive and interdisciplinary foundation for the ethical utilization of AI by learners is synthesized. The objective is to provide a balanced and holistic approach to the employment of AI by students in higher education. By adhering to the synthesized guidelines, it is hoped that AI technologies will be employed in ways that advance the noble objectives of education while safeguarding the ethical, social, and human values at its core.

IEEE's Ethically Aligned Design Standards

The mission of the IEEE's Ethically Aligned Design Standards is to provide a global platform for the development of ethical standards and best practices for the design, development, and deployment of Autonomous and Intelligent Systems (A/IS). The initiative seeks to ensure that these systems are designed and deployed in a manner that is socially responsible, respects human dignity, and is consistent with the principles of human rights. IEEE's Ethically Aligned Design Standards was created to provide a working reference tool for technologists and society as a whole to prioritize ethical considerations, benefits to humanity and the natural environment from the use of A/IS, and mitigate risks and negative impacts, including misuse, as A/IS evolve as socio-technical systems. The five general principles outlined in the IEEE's Ethically Aligned Design Standards are:

1. Human Rights: A/IS should be designed and operated in a way that both respects and fulfills human rights. In practice, this means that AI systems should be developed with a focus on ensuring equitable access, avoiding potential harm or discrimination, and supporting the rights to privacy, freedom of expression, and other fundamental human rights.

2. Human Dignity: A/IS should be designed and operated in a way that respects human dignity, autonomy, and privacy. In other words, AI systems should not degrade, devalue, or objectify humans. They should empower individuals, support their autonomy, and promote their well-being, while ensuring that privacy is maintained and personal data is protected.
3. Fairness: A/IS should be designed and operated in a way that is fair and just. Fairness in AI systems involves treating all users and stakeholders equitably, without discrimination or bias.
4. Non-Maleficence: A/IS should be designed and operated in a way that avoids harm. This principle emphasizes the need for AI systems to prevent harm to individuals, society, and the environment. This includes considering the psychological, social, and economic implications of AI deployment, and ensuring that AI systems are used in ways that align with ethical principles and societal values.
5. Responsibility and Accountability: A/IS should be designed and operated in a way that is transparent and accountable. This principle highlights the need for all stakeholders involved in the AI lifecycle – developers, operators and policymakers – to take responsibility for the ethical implications of AI systems.

Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI

Fjeld et al.'s white paper provides a description of the state of the AI principles field in 2020 by assembling a dataset of thirty-six documents using a purposive sampling method. The dataset includes a range of principle types, from high-level and abstract statements of values to more narrowly focused technical and policy recommendations. The goal of the paper is to facilitate side-by-side comparisons of individual documents and to represent a diversity of viewpoints in terms of stakeholders, content, geography, date, and more. The comparison yielded eight key thematic trends:

1. Privacy: This theme refers to principles that protect individuals' personal data from misuse or unauthorized access. These principles are consent, the ability to restrict processing, the right to erasure, the recommendation of data protection laws, control over the use of data, the right to rectification, and privacy by design.
2. Accountability: This theme refers to the mechanisms through which accountability should be achieved across the lifecycle of an AI system. There are three essential stages: design (pre-deployment), monitoring (during deployment), and redress (after harm has occurred).
3. Safety and Security: This theme's principles include safety, security, security by design, and predictability.
4. Transparency and Explainability: This theme refers to the need for AI systems to be designed and implemented in such a way that their operations can be monitored and understood. This includes the ability to be notified when interacting with an AI, the right to information, open-source data and algorithms, notification when AI makes a decision about an individual, and regular reporting.

5. Fairness and Non-Discrimination: This theme refers to the principles of non-discrimination and the prevention of bias, representative and high-quality data, fairness, equality, inclusiveness in impact, and inclusiveness in design.
6. Human Control of Technology: This theme refers to the idea that humans should have control over technology, particularly AI systems, in order to ensure safety, security, transparency, explainability, fairness, non-discrimination, and the promotion of human values.
7. Professional Responsibility: This theme refers to five principles: accuracy, responsible design, consideration of long-term effects, multistakeholder collaboration, and scientific integrity.
8. Promotion of Human Values: This theme refers to the development and use of AI with reference to prevailing social norms, core cultural beliefs, and humanity's best interests. This includes principles such as human dignity, integrity, freedom, privacy, cultural and gender diversity, fundamental human rights, and protecting and improving our planet's ecosystems and biodiversity.

AI Terms of Use Principles for Higher Education

Incorporating the principles from IEEE's Ethically Aligned Design Standards and Fjeld et al.'s (2020) paper, the following set of AI terms of use principles for students in a higher education setting can be delineated and used in course outlines. The following numbered list includes the principles and an example of violation.

1. Respect Human Rights and Dignity: Students must use AI tools in a manner that upholds and respects human values, including privacy, dignity, and human rights, prioritizing the welfare of all individuals involved in the educational process. A student using an AI-powered sentiment analysis tool to scrutinize the emotions conveyed in their classmates' online discussion contributions and then publicly ridiculing or belittling their peers based on the emotions detected would violate this principle.
2. Ensure Fairness and Non-Discrimination: Students should strive to use AI systems that are fair, unbiased, and non-discriminatory, promoting equal opportunities for all and avoiding perpetuation of existing biases or stereotypes. A student who intentionally uses AI-powered tools with a known history of bias in analyzing content to support an argument for an assignment would violate this principle.
3. Prioritize Privacy and Data Protection: Students must handle personal data responsibly, adhering to relevant data protection regulations and best practices, and ensuring that AI systems used in the educational context prioritize user privacy. A student using an AI-based data analysis tool to collect sensitive information about their peers without obtaining proper consent or following data protection regulations would violate this principle.
4. Promote Transparency and Explainability: Students should seek to understand the functioning, decision-making processes, and underlying algorithms of AI systems used in their coursework, and be prepared to explain their use of AI in their academic work. In a data science project, a group of students use a "Black Box" AI system to analyze a

dataset and generate accurate predictions. However, they cannot explain the system's decision-making process or the parameters they adjusted when questioned during their final presentation. Additionally, they do not disclose their use of this complex AI algorithm. This lack of understanding and transparency about the AI system's operation and their use of AI in their work violates the principle this principle.

5. Commit to Safety and Security: Students must use AI tools in a safe and secure manner, avoiding actions that could compromise the security or reliability of AI systems or cause harm to others. A student sharing their login credentials for an AI-powered educational platform with their classmates, allowing them to access restricted materials or complete assignments on their behalf, thereby manipulating AI adaptive systems, would violate this principle.
6. Maintain Human Control of Technology: Students should use AI systems as a support tool for their learning, ensuring that human agency and decision-making remain at the forefront of their educational experience. A student who excessively relies on an AI-powered essay writing tool to generate their academic work, rather than using it as a supplementary resource to enhance their learning, would violate this principle. For instance, suppose a student is tasked with writing a research paper on an environmental science topic. A violation of this principle would occur if the student simply inputs the topic into the AI tool and lets it generate the entire essay, relying solely on the AI's output. In doing so, the student negates their role in the learning process, excessively relying on technology, and fails to maintain human control in their educational experience.
7. Demonstrate Professional Responsibility: Students must use AI systems ethically and responsibly, considering the potential long-term effects of AI deployment in their academic work, and in collaborating with peers, instructors, and other stakeholders to address potential ethical concerns. A student on a university research team develops an AI tool to predict student success based on various data points. They release the tool without proper oversight or considering potential ethical concerns. Over time, this results in privacy breaches due to misuse of personal data, potential biases and discrimination embedded in the AI's predictions, and an over-reliance on AI in decision-making processes that stifle student growth. This shows a clear failure to demonstrate professional responsibility in the ethical use of AI, with serious long-term impacts.
8. Foster the Promotion of Human Values: Students should use AI systems in a way that aligns with societal values and contributes positively to human flourishing, considering the broader societal context and potential long-term consequences of AI deployment in education. A student using an AI-powered text generator to create fake news articles or social media posts that incite fear, promote disinformation, or provoke animosity between different groups would be violating this principle.

By adhering to these AI terms of use principles, students can ensure that their use of AI systems in higher education is aligned with ethical guidelines that promote a responsible and positive learning experience for all involved.

References

Chaudhry, M. A., & Kazim, E. (2022). Artificial intelligence in education (AIEd): A high-level academic and industry note 2021. *AI and Ethics*, *2*(1), 157–165. https://doi.org/10.1007/s43681-021-00074-z

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI* (SSRN Scholarly Paper 3518482). https://doi.org/10.2139/ssrn.3518482

Lim, J., Seo, J., & Kwon, H. (2022). The Role of Higher Education for the Ethical AI Society. *The International FLAIRS Conference Proceedings*, *35*. https://doi.org/10.32473/flairs.v35i.130609

The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence. (2018). *Ethically aligned design: A vision for prioritizing wellbeing with artificial intelligence and autonomous systems, Version 2*. https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf

Yu, L., & Yu, Z. (2023). Qualitative and quantitative analyses of artificial intelligence ethics in education using VOSviewer and CitNetExplorer. *Frontiers in Psychology*, *14*. https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1061778

Yueh, H.-P., & Chiang, F.-K. (2020). AI and robotics in reshaping the dynamics of learning. *British Journal of Educational Technology*, *51*(5), 1804–1807. https://doi.org/10.1111/bjet.13017