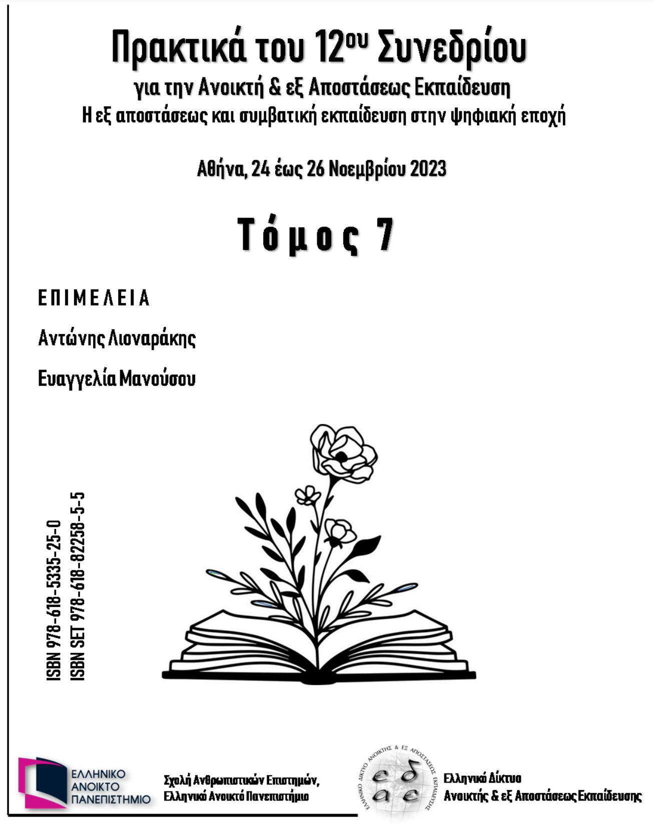
International Conference in Open and Distance Learning

Vol 12, No 7 (2023)

ICODL2023



AI in education: Towards an autotuned magic flute?

Ροδάνθη Τσώνη, Vassilios S. Verykios

doi: [10.12681/icodl.6013](https://doi.org/10.12681/icodl.6013)

AI in education: Towards an autotuned magic flute?

Τεχνητή Νοημοσύνη στην Εκπαίδευση: Ο μαγικός αυλός του μέλλοντος;

Rozita Tsoni
Hellenic Open University
rozita.tsoni@ac.eap.gr

Vassilios S. Verykios
Hellenic Open University
verykios@eap.gr

Abstract

The AI hype has brought on the spotlight technologies that led some circles to argue that we are unleashing AI as a completely untested technology with possibly dangerous implications for society. Despite the scaremongering, AI applications have improved our lives in many fields. In education its implementation has already been made but with minor impact evaluation. This paper examines the concept of AI fairness, focusing on the field of education. In the first section, we present some basic information and definitions of the important concepts of the field of Artificial Intelligence. The next section concerns the application of AI, mainly focusing on educational contexts and how they can enhance the process of teaching and learning. The problem statement section follows, describing several cases where AI produced biased decisions. The next section presents the notion of fairness and equity and how they can form the future of AI for the benefit of humanity. The paper closes with the conclusion section where the main takeaways are presented.

Keywords

Artificial Intelligence (AI), Fairness, Equity, Education, Online Learning

Περίληψη

Η ραγδαία εξάπλωση εφαρμογών Τεχνητής Νοημοσύνης (ΤΝ) έφερε στο προσκήνιο συζητήσεις και προβληματισμούς κατά πόσον η χρήση της συνιστά μια πειραματική εφαρμογή τεχνολογιών με πιθανά επικίνδυνες επιπτώσεις στην κοινωνία. Παρά την κινδυνολογία, οι εφαρμογές ΤΝ έχουν επιφέρει βελτιώσεις σε πολλούς τομείς της ζωή μας. Στην εκπαίδευση ήδη αξιοποιούνται εφαρμογές της ΤΝ χωρίς όμως να έχουν αξιολογηθεί σοβαρά οι πιθανές επιπτώσεις και ο αντίκτυπός τους. Στην παρούσα εργασία εξετάζουμε την έννοια της δικαιοσύνης στην ΤΝ, εστιάζοντας κυρίως στον τομέα της εκπαίδευσης. Στην πρώτη ενότητα παρουσιάζουμε μερικές βασικές πληροφορίες και ορισμούς για τις σημαντικές έννοιες του τομέα της ΤΝ. Η επόμενη ενότητα αφορά την εφαρμογή της ΤΝ με εστίαση στην εκπαίδευση και το πώς μπορεί να βελτιώσει τη διαδικασία διδασκαλίας και μάθησης. Ακολούθως, περιγράφονται αρκετές περιπτώσεις όπου η τεχνητή νοημοσύνη παρήγαγε μεροληπτικές αποφάσεις ώστε να αναδειχθούν οι κίνδυνοι από τη λανθασμένη χρήση της. Η επόμενη ενότητα παρουσιάζει την έννοια της δικαιοσύνης και της ισότητας και πώς μπορούν να διαμορφώσουν το μέλλον της τεχνητής νοημοσύνης προς όφελος της ανθρωπότητας. Τέλος συζητούνται τα βασικά συμπεράσματα.

Λέξεις-κλειδιά
Τεχνητή Νοημοσύνη (ΤΝ), Δικαιοσύνη, Ισότητα, Εκπαίδευση, Διαδικτυακή Μάθηση

Introduction

In the Middle Ages legend of the Brothers Grimm called Pied Piper of Hamelin, the mayor of a city invited a piper to save them from a massive mise invasion. The piper used his flute to attract the mice and keep the city clean. Upon the refusal of the mayor to pay him, the piper played a different tune attracting this time, almost all the children of the city, causing unsufferable pain to all their parents, as an act of revenge. We live in a highly interesting era where technological advances are so fast that exceed our capacity to fully understand their possibilities, dangers and future impact. However, AI is present, widely used, surely inevitably impacting our lives. Our experience until now shows that AI learns by itself but it learns from humans. Therefore, it is not flawless. In this work, we approach some of the main challenges concerning fairness in AI in general and its application in educational settings. If an AI application acts as a magic flute, who plays the melody? What are the consequences? What if this flute is autotuned and how do stakeholders should act to ensure best practices for future generations?

The AI hype has brought on the spotlight technologies that led some circles to argue that we are unleashing AI as a completely untested technology with possibly dangerous implications on society. However, AI technologies have long been used and tested in various applications. Stephen Downs in his article *"How to Use AI"* (2023) describes several applications such as translating labels via camera shots, weather predictions, metro guides, biometric login systems, bank security systems, adaptive cruise control systems, grammar checking applications, content recommendation in movies and music apps, automatic text generators. Yet, researchers stress that AI inherently magnifies latent characteristics present in its initial data, thereby reinforcing its inherent assumptions. Algorithms are mostly trained on data imbued with human bias that they reproduce and sometimes intensify its impact. This poses a significant challenge, especially as long as there is a widespread belief in the community that algorithms operate impartially. Therefore, ethical issues concerning fairness in AI are in the spotlight, particularly in the field of education which plays a crucial role in shaping the future society.

In the first section, we present some basic information and definitions of the important concepts of the field of Artificial Intelligence. The next section concerns the application of AI focusing on educational context and how they can enhance the process of teaching and learning. The problem statement section follows, describing several cases where AI produced biased decisions. The next section presents the notion of fairness and equity and how they can form the future of AI for the benefit of humanity. The paper closes with the conclusion section where the main takeaways.

Definitions and background knowledge

In this section basic concepts including Artificial Intelligence and Machine Learning will be defined and explained.

Artificial Intelligence: Artificial Intelligence (AI) is a term that was introduced back in 1955 by Stanford Professor John McCarthy as "*the science and engineering of making intelligent machines".* Now it constitutes an area of study in the field of computer science. AI broadly concerns the capacity of computers to engage in human-like thought processes such as learning, reasoning, and self-correction. A more concrete definition was given by McCarthy, J. (2004): *"It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable."*

Machine Learning: Machine Learning (ML) is the part of AI studying how computer agents can improve their perception, knowledge, thinking, or actions based on experience or data. The dataset used to train the ML model is called *training dataset*. ML draws from computer science, statistics, psychology, neuroscience, economics and control theory. It can be divided into three broad categories: Supervised Learning, Unsupervised Learning and Reinforcement Learning. In supervised learning, a computer learns to predict human-given labels, such as dog breed based on labeled dog pictures Unsupervised learning does not require labels. Usually it performs prediction tasks such as trying to predict each successive word in a sentence. Reinforcement learning lets an agent learn action sequences that optimize its total rewards, such as winning games, without explicit examples of good techniques, enabling autonomy.

Deep Learning (DL): Deep Learning is a part of the ML techniques' family (Figure 1) leveraging neural networks of three or more layers:

i. Input layer: This is where data enters the system.
ii. Hidden layers: These layers process and convey data to other layers.
iii. Output layer: The outcome or prediction is generated in the output layer.

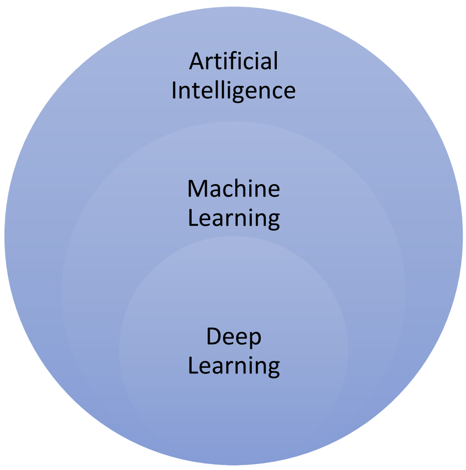


Figure 1: *AI, ML and DL overlap*

Neural networks aim to emulate human learning processes by assimilating and scrutinizing extensive information, often referred to as training data. Through iterative repetition of a given task with this data, neural networks enhance their accuracy over time. This iterative learning process is mimicking the process that individuals use to study and practice in order to refine their skills.

Table 1: The differences between ML and DL (Rivas, 2020)

Machine Learning	Deep Learning
A subset of AI	A subset of machine learning
Can train on smaller data sets	Requires large amounts of data
Requires more human intervention to correct and learn	Learns on its own from environment and past mistakes
Shorter training and lower accuracy	Longer training and higher accuracy
Makes simple, linear correlations	Makes non-linear, complex correlations
Can train on a CPU (central processing unit)	Needs a specialized GPU (graphics processing unit) to train

Protected or sensitive attributes: A sensitive attribute refers to a characteristic or feature of an individual that is deemed as potentially sensitive or subject to privacy concerns. In the context of machine learning and data analysis, sensitive attributes are often variables such as race, gender, political orientation, health issues, disability, age, ethnicity, or any other factor that could lead to potential discrimination or bias if not handled appropriately.

Input and output similarity: Input similarity refers to the similar characteristics of the entities of the training dataset. It is a method for quantifying the resemblance between entities in the context of individual fairness, as well as a mechanism for categorizing entities into groups for group fairness. Output similarity basically refers to the similar treatment of the entities either individually or per group, in the results of an algorithm.

Ranking: Ranking, is an ML approach that involves organizing items in a specific order based on their relevance or importance. In the ranking technique, a model is trained to predict the ranking of one item over another by assigning a "score" to each item. Higher-ranked items have higher scores, while lower-ranked items have lower scores. The model utilizes these scores to predict which item ranks higher than the other, directly influencing the order in which items are presented (Rahangdale, & Raut, 2019). Ranking holds significance in various information retrieval domains, such as e-commerce, social networks, and recommendation systems. The effectiveness of recommendation systems depends on their ability to ensure that similar and relevant products are presented in a manner that leads the user to click or make a purchase. In a broad sense, ranking may appear similar to regression models. However, while a basic regression model can estimate the likelihood of a user buying a product, a more practical approach involves the use of ranking techniques. This allows for strategic ordering or prioritization, intending to maximize the likelihood of obtaining a purchase. The prioritization of items significantly influences users' decisions positively. Several applications use ranking, namely: Search engines, recommendation systems for online shopping, music platforms, travel agencies, smart TV platforms, etc.

Recommendation: A recommendation system, also known as a recommender system, falls under the category of information filtering systems designed to offer personalized suggestions to users (Ricci et al., 2021). These suggestions are aimed at helping users make decisions in various contexts, such as selecting a product to buy, choosing music to listen to, or picking online news to read. Recommender systems prove especially valuable when users face the challenge of selecting an item from a vast array of choices offered by a service (Resnick & Varian, 1997). A recommendation system functions as a data filtering engine employing Deep Learning principles and algorithms to propose potential products based on users' past preferences or additional filtering criteria. The underlying idea of these algorithms involves identifying patterns in consumer behavior or similar behaviors related to a particular service or product. The approach to data collection varies significantly depending on the nature of the products or services being offered: an online shop would gather data through review ratings, while platforms like YouTube store the number of likes and dislikes of its videos. The life cycle of a recommendation system consists of seven steps (Hrnjica et al, 2020):

1. Data collection
2. Data storage
3. Data filtering
4. Data analysis
5. Model Evaluation and Testing

6. Model deployment
7. Online Machine Learning

AI in Education

Distance Learning is one of the first field of implementation of AI techniques in education. Currently, we have far passed the era where online and the so-called conventional education were two distinct methods of education. Almost every class is connected and students of all ages in the larger part of the globe have access to remote information. AI promotes personalisation even in large audiences. It curates content and makes choices on behalf of the users. AI is, by all means, changing the educational landscape. There is an extraordinary range of AI approaches in education. Intelligent tutoring systems, are the most common ones. Holmes et al. (2023) describe six main possible applications of AI in education. Namely:

1. Collaborative learning
2. Students' forum monitoring
3. Continuous assessment
4. AI learning companions
5. AI teaching assistant
6. Research Tool to Further the Learning Sciences

In the field of special education, AI can enable the provision of differentiated education for a wide range of people (Trewin, Shari, et al., 2019). Enormous benefits can come with personalized learning (Mossis et al., 2010; Kouvara et al., 2022) that would not be possible in typical classrooms bringing together all peers and eliminating discrimination. A lot of progress has been made towards inclusive education. Yet, integrated learning environments are still unavailable to a large number of people with disabilities (Dudley-Marling & Burns, 2014), especially in K-12 classrooms (Shaheen, & Lazar, 2018). Similar problems have been identified in post-secondary, higher and online education where many platforms present low levels of accessibility (Burgstahler, 2015; Cinquin et al., 2019; Walters, 2022). These issues should shape future targeting and lead the way to a more inclusive and democratic education for the next generations to come.

The application of AI in education has been featured as one of the most pivotal developments of the century (Becker et al., 2018; Seldon with Abidoye, 2018). However, there is certain skepticism about its implementation in the classroom that is rather slow and unorganized resampling the adoption of computers in education when Cuban (2001) characteristically stated: *"oversold yet underused in classrooms".* Very recently a vast invasion of an innovative, beyond-the-ordinary, tool has alerted the educational community. ChatGPT came to shake the foundations of educational design. It can automatically generate natural language and perform a series of tasks such as finding titles, generating and summarizing content, and organizing and presenting ideas. It is now up to the tutors, the educational designers and the educational stakeholders in general to adapt the course delivery, having in mind that students have at their disposal a new powerful tool and leverage in a beneficial way its capacities. However, there is a possibility that ChatGPT technology might become widespread before institutions can adapt their policies. A more effective approach would involve addressing issues stemming from ChatGPT use while carefully considering the platform's potential advantages and disadvantages. When groundbreaking educational technology becomes accessible to the public, it becomes the responsibility of educators and policymakers to tackle rising challenges and devise strategies to eliminate inefficient educational practices (Baidoo-Anu, & Ansah, 2023; Mhlanga, 2023).

Not everything in the garden is rosy: the problem statement

A way to study the acceptance of a new situation, especially technology-related is the Hype Circle (Figure 1). People's attitudes follow this pattern in many innovative circumstances. The hype cycle serves as a visual representation that illustrates the evolution, acceptance, and real-world implementation of specific technologies. The hype cycle aids in evaluating emerging trends across various technologies (Predana et al., 2021). Its applicability is particularly notable in navigating and evaluating the diverse activities encompassed within the realm of AI (Neuhofer et al., 2021). This phenomenon aligns with Amara's law, emphasizing the tendency to overestimate a technology's impact in the short term while underestimating its long-term effects (Crews 2019).

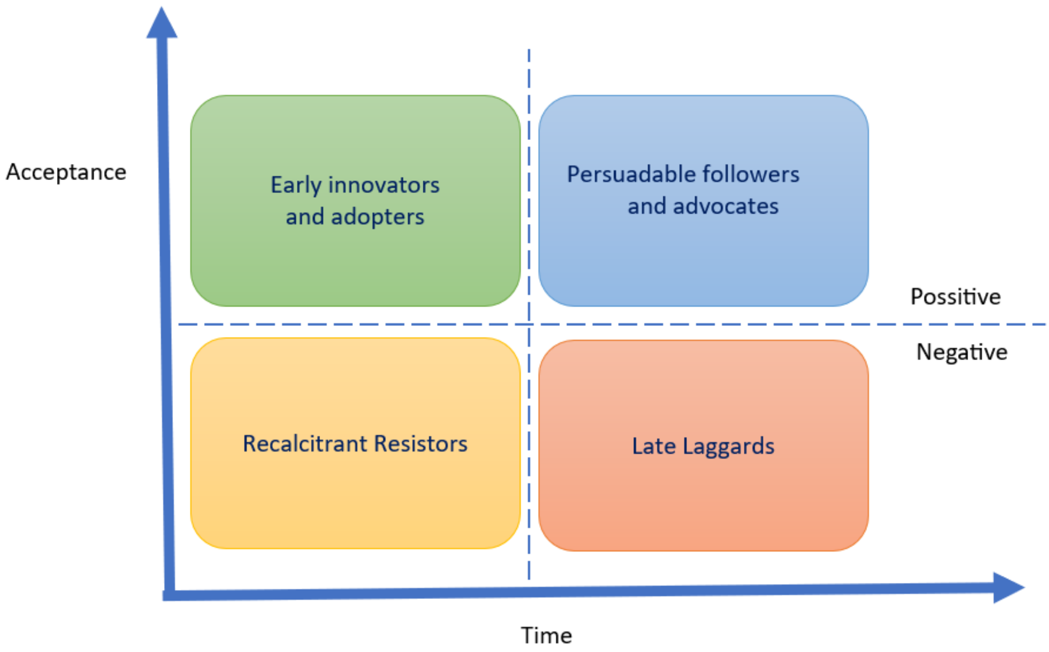


Figure 2: *The Hype Cycle*

Another model widely spread in the business sector for guiding decisions concerning change is the Change Curve (Figure 3). The Change Curve was originally developed by Elisabeth Kübler-Ross (Kübler-Ross model) to describe the five stages of grief and later was applied to unexpected events and unpredictable innovations (Lukianov et al., 2020). However, this is not the case in AI implementation. Despite its vast innovative potential and its large impact on our lives, neither the Hype Circle nor the Change Curve are representative of AI public acceptance. This is due to a significant difference between other innovations and AI invasion. Both abovementioned models require awareness of usage, while AI applications come to our lives silently, facilitating our transactions without even noticing their existence.

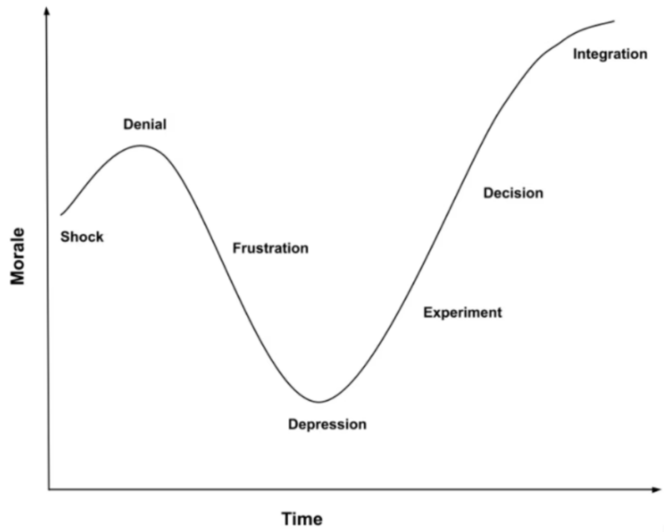


Figure 3: *The Change Model*

AI is most of the time the result of an assembly of the footprint of human action containing biases, prejudices and misconceptions. The quality of the data that train a ML algorithm has a decisive impact on the produced application.

Numerous research papers highlight the potential of AI-based decision support systems to unintentionally encode human biases and even introduce new ones (Chouldechova, & Roth, 2020). Additionally, it seems that there is an inherent tendency of people to approve of such biases. The study of Kay et al. (2015) showed that people tend to rate higher search results when these results are consistent with stereotypes. Many biased results have been found in the relevant literature. Female artists are not given equal exposure in music recommendations on popular music platforms (Ferraro, 2021). Names used more often by individuals of color, regardless of gender, are significantly more prone to trigger advertisements related to arrest records (Sweeney, 2013). Additionally, when employing a tool known as Adfisher, it was observed that configuring the gender to "female" led to receiving advertisements predominantly geared towards lower-paying job opportunities. Furthermore, in the context of word embeddings, it was noted that the vector representing "computer programming" is more closely associated with men than with women.

Ali et al., in their work, showed that group targeting may inadvertently influence ad delivery, causing certain users to be less likely to see specific ads based on their demographic characteristics. It was revealed that such skewed delivery occurs on

Facebook, influenced by market dynamics, financial optimization, and the platform's predictions regarding ad "relevance" to different user groups. Both the advertiser's budget and ad content significantly contribute to Facebook's ad delivery skew, particularly along gender and racial lines for "real" ads related to employment and housing opportunities, even when targeting parameters are set to be highly inclusive.

The COMPAS system, a commercially employed tool utilizing a risk assessment algorithm to predict certain future crime categories, is utilized in U.S. courts to aid in bail and sentencing determinations. Research (Angwin, et al., 2016) revealed that the false positive rate, indicating individuals labeled as high risk by the tool who did not re-offend, was nearly twice as high for African-American defendants compared to their white counterparts.

The abovementioned paradigms can justify the concerns of society about AI fairness issues. A long and extensive discussion has begun about the ethical impact of AI usage. All the aspects of our lives are affected in a way by applications that people use, many times even without being aware of it. ChatGPT (OpenAI, 2023) summarizes the danger of its misuse: *"The main danger of AI models like ChatGPT for humanity is the potential for misuse and unintended consequences. Some of the key concerns include:*

1. *Misinformation and Disinformation: AI models can generate false or misleading information, making it easier for bad actors to spread disinformation and manipulate public opinion. This can have serious implications for democracy and societal stability.*
2. *Bias and Discrimination: AI models can inherit biases present in the training data, leading to discriminatory or unfair outcomes. This can reinforce existing societal biases and inequalities.*
3. *Privacy Concerns: AI models can generate highly personalized content, potentially infringing on individuals' privacy by revealing sensitive information or creating deepfakes.*
4. *Job Displacement: Automation driven by AI could lead to job displacement in various industries, potentially causing economic and social disruption.*

5. *Security Risks: AI can be used by malicious actors for cyberattacks, fraud, and other illicit activities, making it a cybersecurity concern.*
6. *Ethical and Moral Dilemmas: The use of AI in various applications, such as autonomous weapons or decision-making in critical areas like healthcare, raises ethical and moral questions.*
7. *Loss of Human Expertise: Overreliance on AI for decision-making may lead to a decline in human expertise and critical thinking skills.*
8. *Lack of Accountability: AI decisions can be difficult to explain or hold accountable, which can pose challenges in ensuring responsible and transparent AI systems.*
9. *Superintelligent AI: As AI technology advances, there are concerns about the long-term risks associated with developing superintelligent AI systems that may surpass human capabilities and potentially act against human interests."*

About Fairness

Fairness in general, means the lack of discrimination bias. The new regulations establish obligations for both providers and users based on the level of risk posed by artificial intelligence. While many AI systems involve minimal risk, they have to be assessed. Nguyen et al., 2023 conducted a thematic analysis of the guidelines about AI in education ethics derived from international organization reports such as UNESCO Ethics AI (AHEG, 2020), UNESCO Education & AI (Miao et al., 2021), Beijing Consensus (UNESCO, 2019), OECD (Organization for Economic Co-operation and Development, 2021), European Commission (2019), and European Parliament Report AI Education (2021). The result was to produce a condensed collection of seven principles, namely:

1. Principle of governance and stewardship
2. Principle of transparency and accountability
3. Principle of sustainability and proportionality
4. Principle of Privacy
5. Principle of Security and Safety
6. Principle of inclusiveness
7. Principle of human-centered AIED

Fairness, equality and Equity

The sensitive nature of fairness demands a deep understanding of the concepts related to it. It would be expected that a blind process while handling data would ensure equal treatment. However, there are quasi-features that can skew a seemingly fair process and produce biased results. From a sociological point of view, it is important to underline the difference between equity and equality. While equality simply states that different entities would be treated equally, equity demands that entities should be treated according to their needs, ensuring similar results even in disadvantaged groups.

Even, simple, seemingly insignificant applications can contain bias. For example, in a smart TV platform where there is a large number of options, viewers' choices are mainly defined by the recommendation systems. These systems base their decision on what was viewed before (a decision that contains problems of embedded bias), leading to a suggestion loop that lowers the diversity and amplifies bias. To understand and try to reduce the effects of such phenomena we have to thoroughly examine the concept of fairness, considering multiple aspects of the problem.

The concept of fairness when it comes to be implemented on AI applications can be taxonomized according to *level*, *side* and *outport multiplicity* (Pitoura et al., 2021).

The *level* distinguishes between individual and group fairness. Individual fairness means that similar entities should be treated equally, while in group fairness, different groups of people according to a sensitive attribute should be treated equally. Input similarity, output similarity.

To classify fairness according to the *side*, the producer and consumer classes are used. It mainly concerns recommendation systems; thus, distinguishing sides ensures that the rights of both participants in a transaction involving the recommendation of products and goods are protected. As was earlier discussed, recommendation systems, extensively used across various domains such as movies, jobs, and courses, often encounter unfairness in predictions. As predictions rely on observed data, they may inadvertently inherit pre-existing biases. To address this concern, (Yao, & Huang, 2017) introduce consumer-side unfairness metrics, which examine the variation in prediction behavior between protected and non-protected users. On the producer side, fairness may be quantified in terms of two distinct types of benefit functions: exposure and relevance (Gómez et al., 2022). An ideal system would balance both types of fairness to provide fairness guarantees for both sides.

When we seek to make a purchase or choose something to use there are also two sides involved the side of the user and the side of the item that is being ranked. Therefore, user-side fairness concentrates on the individuals who access or consume the data items in a ranking, such as a search result or recommendation. In broad terms, the aim is to have similar users or user groups receiving comparable rankings or recommendations. For instance, if gender serves as the protected attribute for a user receiving job recommendations, the expectation is that the user's gender does not impact the job recommendations they receive. On the other hand, the item-side fairness directs attention to the items undergoing ranking or recommendation. In this context, the objective is to ensure that comparable items or categories of items are ranked or recommended similarly, maintaining a consistent position in a ranking. This represents the primary form of fairness discussed thus far. For example, if we designate political orientation as the protected attribute for an article, we seek to prevent this attribute from influencing the article's ranking in a search result or news feed.

Lastly, based on the *output multiplicity*, there is a distinction between single output and multiple output fairness. In the context of multiple output fairness, we seek eventual or amortized fairness for consumers or producers. This implies that, over a sequence of rankings or recommendations, consumers or producers should be treated fairly as a collective, even if there are instances of unfair treatment in one or more individual rankings or recommendations within the sequence.

Methods to ensure fairness in ranking

AI scientists are trying to implement methods to ensure that AI applications minimize bias and deliver fair results. These methods can be applied in three layers: the pre-processing approaches, the in-processing approaches and the post-processing approaches. Broadly, the pre-processing approaches focus on removing bias and discrimination from the training data and evaluating the existence of underrepresented groups. The in-processing methods concern the data analysis and processing techniques, modifying the existing algorithms or introducing new ones in order to produce fair results. The post-processing approaches are applied in the results of an ML process ensuring output similarity modifying -when needed the results- to remove discrimination or unequal treatment.

Barriers and Challenges

The awareness of the risks of injustice in AI is the first step towards a solution. The AI case is a multidisciplinary challenge that demands the collaboration of scientists of a very wide spectrum and background. There is a long way to go and the next steps have to be made rapidly in order to effectively follow the technological evolution. According to Pitoura et al., (2021), there are eight open challenges in ensuring fairness in AI systems:

i. A codification of definitions
ii. Lack of data
iii. A unified approach for the data pipeline
iv. Lack of evaluation tools
v. Lack of real applications of fairness
vi. A multi-level architecture of value systems and algorithms
vii. Relating algorithmic fairness with other notions of fairness in systems.
viii. Fairness in other domains

We live in the hype of virtual agents and bots. There is more of an issue of digital natives or immigrants. Children are living in a different world from their parents and teachers concerning technological use. Therefore, the use of AI in education requires urgent attention. Students are used to operating applications that depend on huge amounts of personal data and efficient algorithms, raising privacy and ethical issues. Thus, educational institutes should be role models of ethical AI usage, leading the way for other organisations.

Additionally, educational stakeholders should also consider the black box effect. Accordingly, the internal operations of an algorithm remain hidden from the user, while the input data and corresponding output are known, the code or logic responsible for generating this output is not accessible for examination. In an educational setting, a method can be rejected based on the “average” while it might be more effective for particular individuals or groups. Thus, justice in AI is an upcoming challenge of high complexity and huge importance.

Fenu et al. (2022) organize the challenges of using AI in education into four broad categories. The first refers to *the "Legacies of educational systems of oppression".* Educational AI technologies are forged under certain circumstances. Hence, there are not only accessibility issues that have to be solved in order to have fair representation in training data for decision making, but also, we have to carefully consider the socio-economical settings and educational practices that shape the existing conditions. The second challenge is the *"Biopolitical educational technologies",* that is the challenge to keep education clear from behavioural management based on ideologies that reproduce structural injustices. Thirdly, there is widespread skepticism about *"surveillance technologies".* The use of AI simply for surveillance is by definition unethical. Multimodal data (image, eye movement recording, biometric characteristics, voice, etc.) demand huge attention regarding their collection, storage, analysis and usage. The fourth challenge concerns *"at-risk prediction technologies".* Usually, they are used to label students. However, the labelling should aim strictly to prevent a future failure. Additionally, there might be issues of justice when, for example, some students attract more attention than their peers simply because they are not working as hard as they should.

Unlike a common feeling amongst the educational circles, teachers are not really threatened by AI, like conventional education was never really threatened by Distance Education. According to Harari, (2023), for AI to become threatening at least three important benchmarks would have to be reached: to attain consciousness, to attain emotion and to be able to navigate to the real world. The implementation of AI in education is not aiming to replace emotional, multileveled relations between tutors and their students. AI can support and enhance this process. Instead of providing decision-making, AI applications can deliver decision support. A common misunderstanding is to see AI systems as an expert's opinion when they rather express the wisdom of the crowds summarising what an average person similar to us would do or choose.

Towards an attuned magic flute? Conclusions

In the legend people of Hamelin paid their unwillingness to spend their money with the loss of their children. What will be the price of our own unwillingness to seriously consider the conditions under which AI would be embedded in our lives? Technology itself is not inherently dangerous. However, its misuse or lack of responsible oversight can pose significant risks. To mitigate AI risks, it's essential to develop and implement responsible AI practices, including transparent AI development, bias mitigation, robust cybersecurity measures, and ethical guidelines. Additionally, ongoing research and regulatory efforts are crucial to ensure that AI technologies benefit humanity while minimizing potential harm.

One important feature is that AI applications are gaining the ability to develop deep and intimate relationships with humans (Bengio et al., 2023). AI systems are self-improved to the point where even their creators do not know their full capacities in advance. Their ability to manipulate and generate language at a level that surpasses the average human ability. The implications in education and the wider society are yet to be discovered. Human rights are not a biological reality. Therefore, we constantly have to focus our efforts on preserving justice at all levels of our interactions whether they involve humans, or this interaction also involves technological applications. These efforts require awareness and knowledge. Thus, issues as broad as accuracy, choice, predictions, privacy, fairness, and ethics are what we should be teaching school and university students, structuring a society of well-inform, conscious and active citizens.

References

Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A., & Rieke, A. (2019). Discrimination through optimization: How Facebook's Ad delivery can lead to biased outcomes. *Proceedings of the ACM on human-computer interaction*, 3(CSCW), 1-30.

Angwin J., Larson J., Mattu S. & Kirchner L. (2016). Machine bias. *ProPublica*. Retrieved by: https://www.propublica.org/article/machine-bias-risk-assessments-incriminal-sentencing

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62.

Bengio, Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Harari, Y. N., ... & Mindermann, S. (2023). *Managing AI Risks in an Era of Rapid Progress*. arXiv preprint arXiv:2310.17688.

Burgstahler, S. (2015). Opening doors or slamming them shut? Online learning practices and students with disabilities. *Social Inclusion*, 3(6), 69-79.

Chouldechova, A., & Roth, A. (2020). A group of industry, academic, and government experts convene in Philadelphia to explore the roots of algorithmic bias. *Communications of the ACM*, 63(5), 82-89.

Cinquin, P. A., Guitton, P., & Sauzéon, H. (2019). Online e-learning and cognitive disabilities: A systematic review. *Computers & Education*, 130, 152-167.

Crews, C. (2019). How Far to Look Ahead?. *Research-Technology Management*, 62(5), 59-62.

Cuban, L. (2001). *Oversold and underused: Computers in the classroom*. Harvard university press.

Downes S. (2023). How to use AI. *Online Educa Berlin 2023*. Retrieved by: https://www.downes.ca/cgi-bin/page.cgi?post=75058

Dudley-Marling, C., & Burns, M. B. (2014). Two perspectives on inclusion in the United States. *Global Education Review*, 1(1), 14-31.

Fenu, G., Galici, R., & Marras, M. (2022). Experts' view on challenges and needs for fairness in artificial intelligence for education. In *International Conference on Artificial Intelligence in Education* (pp. 243-255). Cham: Springer International Publishing.

Ferraro, A., Serra, X., & Bauer, C. (2021). Break the loop: Gender imbalance in music recommenders. In *Proceedings of the 2021 conference on human information interaction and retrieval* (pp. 249-254).

Gómez, E., Boratto, L., & Salamó, M. (2022). Provider fairness across continents in collaborative recommender systems. *Information Processing & Management*, 59(1), 102719.

Harari Y. N., 2023. AI and the future of humanity. *Yuval Noah Harari at the Frontiers Forum*. Retrieved by: https://www.youtube.com/watch?v=LWiM-LuRe6w

Holmes, W., Bialik, M., & Fadel, C. (2023). *Artificial intelligence in education*. Globethics Publications.

Hrnjica, B., Music, D., & Softic, S. (2020). Model-based recommender systems. *Trends in Cloud-based IoT*, 125-146.

Kay, M., Matuszek, C., & Munson, S. A. (2015, April). Unequal representation and gender stereotypes in image search results for occupations. In Proceedings of *the 33rd annual acm conference on human factors in computing systems* (pp. 3819-3828).

Kok, J. N., Boers, E. J., Kosters, W. A., Van der Putten, P., & Poel, M. (2009). Artificial intelligence: definition, trends, techniques, and cases. *Artificial intelligence*, 1, 270-299.

Kouvara T., Tsoni R., Verykios V. S. (2022). Openness Is not Enough Without Inclusiveness: The Learning Analytics Parable. In *the proceedings of I·HE2022*, EADTU

Lukianov, D., Kolesnikova, K., Mezentseva, O., & Rudenko, V. (2020). The kübler-ross factor in managing the performance of technical and socio-economic systems. *Scientific Journal of Astana IT University*, 32-43.

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning (February 11, 2023).

Morris, R. R., Kirschbaum, C. R., & Picard, R. W. (2010, October). Broadening accessibility through special interests: a new approach for software customization. In Proceedings of *the 12th international ACM SIGACCESS conference on Computers and accessibility* (pp. 171-178).

Neuhofer, B., Magnus, B., & Celuch, K. (2021). The impact of artificial intelligence on event experiences: a scenario technique approach. *Electronic Markets*, 31, 601-617.

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, 28(4), 4221-4241.

OpenAI. (2023). ChatGPT [Large language model]. https://chat.openai.com

Perdana, A., Robb, A., Balachandran, V., & Rohde, F. (2021). Distributed ledger technology: Its evolutionary path and the road ahead. *Information & Management*, 58(3), 103316.

Rahangdale, A., & Raut, S. (2019). Machine learning methods for ranking. *International Journal of Software Engineering and Knowledge Engineering*, 29(06), 729-761.

Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, 40(3), 56-58.

Ricci, F., Rokach, L., & Shapira, B. (2021). Recommender systems: Techniques, applications, and challenges. *Recommender Systems Handbook*, 1-35.

Rivas, P. (2020). *Deep Learning for Beginners: A beginner's guide to getting up and running with deep learning from scratch using Python*. Packt Publishing Ltd.

Shaheen, N. L., & Lazar, J. (2018). K–12 technology accessibility: The message from state governments. *Journal of Special Education Technology*, 33(2), 83-97.

Sweeney, L. (2013). Discrimination in online ad delivery: Google ads, black names and white names, racial discrimination, and click advertising. *Queue*, 11(3), 10-29.

Trewin, S., Basson, S., Muller, M., Branham, S., Treviranus, J., Gruen, D., ... & Manser, E. (2019). Considerations for AI fairness for people with disabilities. *AI Matters*, 5(3), 40-63.

Walters, A. P. (2022). *Evaluating the Effectiveness of Online Faculty Development in Creating Accessible Content* (Doctoral dissertation, The University of Toledo).

Yao, S., & Huang, B. (2017). Beyond parity: Fairness objectives for collaborative filtering. *Advances in neural information processing systems*, 30.