Artificial Intelligence Ethics from the Perspective of Educational Technology Companies and Schools

Kousa, Päivi

Artificial Intelligence Ethics from the Perspective of Educational Technology Companies and Schools

Päivi Kousa and Hannele Niemi

Contents

1 Introduction

In this chapter, AI will be presented in contemporary educational contexts. The aim is to understand what kind of ethical challenges EdTech companies and schools have and how those challenges affect their daily work. As technology evolves at an accelerating pace and the education sector seeks to keep up, rapid actions are needed to avoid the ever-growing gap between EdTech companies and schools. First, companies' and schools' reflections during interviews are presented inductively with their own concepts based on two Finnish case studies. Thereafter their thoughts are contextualised in terms of five ethical principles by Morley et al. (2020).

Artificial intelligence (AI) has become part of the global discussion and our everyday lives more than ever, although AI and machine learning have been among us for decades (Turing 1950/2009). AI is influencing almost all levels of our economy and society. For example, it enables people to use new tools and

P. Kousa (✉) · H. Niemi
Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland
e-mail: paivi.m.kousa@jyu.fi; hannele.niemi@helsinki.fi

applications, e.g. transportation, services, healthcare, education, public safety and security, employment and workplace, and entertainment (Stone et al. 2016; Littman et al. 2021). All these changes have fundamental influences on organisations which establish new demands which then need to be fulfilled by their staff developing new competences. New technology and advanced methods in computing with AI applications are increasingly used also in education. Globally, there are several common AI-related practices and tools for education and learning, such as teaching robots, intelligent tutoring systems (ITS), online learning, and learning analytics. Augmented and virtual realities are interactive systems often used for competence training, especially in many areas of life-long learning (e.g. Grover and Pea, 2018).

Although there is a global consensus that AI should be ethical, many problems exist in defining the values embodied in ethical guidelines. Ethical guidelines are not conceptually congruent but are rather open to a wide range of interpretations (e.g. Jobin et al. 2019). Many companies find general guidelines useless and prepare guidelines of their own instead (Hagendorff 2020). Cath (2018) suggests that universities and other organisations (e.g. policymakers and schools) could offer a leading, research-based, and objective role in the development of ethical guidelines since industry-produced guidelines may be too subjective. There are also fears that companies are too involved in drafting legislation and guidelines which serve to pursue their own interests (Cath 2018). However, cooperation is needed since so many parties are involved in the AI ecosystem, such as policymakers, universities, schools, and industries. Yet discussions between developers and researchers have lasted decades without sufficient outcomes (Bostrom and Yudkowsky 2011). Secondly, products and services based on AI are difficult or, in many cases, almost impossible to explain (Goebel et al. 2018), although their explainability and interpretability would enhance the fairness, transparency, and accountability needed for those who use AI products and services (Cath 2018). Thirdly, schools need education and guidelines which can be implemented during their daily work. Nnaji (2019) discusses how ethical conflicts in schools have more to do with how the technology is used than in the technology itself. He states that different applications are simply tools to help students and teachers in their work but should not be blindly trusted or allowed to guide school activities without critical considerations. AI in education presents serious challenges in relation to the issues of student privacy, accuracy, data ownership, accessibility, and integrity which need to be addressed (Nnaji 2019).

2   AI in Education and Learning

The increased use of AI for education and learning has promoted many opportunities as well as major challenges (Torresen 2018). According to United Nations Educational, Scientific, and Cultural Organization, UNESCO (2019), there are six major challenges related to AI in education (AIED): (1) lack of comprehensive public policy on AI, (2) unequal opportunities to use AIED, (3) lack of adequate teacher education, (4) lack of development of quality and inclusive data systems,

(5) lack of significant AI-related research, and (6) lack of ethics and transparency in data collection, use, and dissemination. Concerns with data privacy and ownership issues, and the safety of public/private interfaces, have raised questions especially in educational fields (e.g. Dignum 2018). Many researchers and international organisations claim that AI should be trustworthy—lawful, ethical, and socially as well as technically robust (High-Level Expert Group on Artificial Intelligence, AI HLEG 2019a). In education and learning, ethical challenges have grown in tandem with technological development, as AI trustworthiness has become increasingly important (e.g. Stanford Institute for Human-Centered AI, HAI 2020). Although AI has many benefits for learning, the educational field has faced many challenges in relation to equity, data management, decision-making, and human and machine learning (e.g. Stone et al. 2016). When AI is implemented in educational contexts, education stakeholders must be able to trust that the entire design processes of AI-based solutions are ethical and that the algorithms are designed in accordance with ethical principles that suit the values of the school world.

Yet Holmes et al. (2021) emphasise that ethics is not a straightforward concept in the context of education. They urge distinguishing between 'doing ethical things' and 'doing things in an ethical way'. They suggest that AIED technologies should include specific 'ingredient lists' like in food or medicine products. This proclamation in labelling would increase the understandability and transparency of the AI-based solutions. In practice, this could mean that the user (e.g. a teacher or a student) would be informed of the limitations or benefits of the product beforehand. Goebel et al. (2018) remind us that efforts have been made to explain complex AI systems for decades. It can be concluded that many ethical challenges are present when designing AI-based tools and services for education. In addition, ethical factors are always present in education product design (e.g. in schools and workplaces), since the purpose is to exert influences on peoples' minds, behaviours, and lives. This pervasive influence of education makes educational AI solutions even more challenging to develop. Although AI can provide many beneficial solutions to existing educational challenges, there are many new problems that need to be solved between EdTech companies and schools who use the solutions that are developed. The recent coronavirus disease 2019 (COVID-19) increased distance learning and thus the urgent need for teachers and students to use digital applications and understand how they work (e.g. Niemi and Kousa 2020).

3   Many Ethical Guidelines and Principles for AI

Numerous international, national, governmental, organisational, and company-based guidelines exist for ethical AI. For example, the European Commission's high-level group on artificial intelligence (AI HLEG) has published four deliverables: ethics guidelines for trustworthy AI with 7 key requirements (AI HLEG 2019a), policy and investment recommendations for trustworthy AI with 33 recommendations (AI HLEG 2019b), assessment list for trustworthy AI which can be

used as a practical aid when implementing requirement into practice (AI HLEG 2020a), and sectoral considerations on the policy and investment recommendations which provide examples concerning how and where regulations can be implemented (AI HLEG 2020b). The guidelines are developed in collaboration with an AI alliance including 4000 stakeholders (e.g. European Union/EU citizens, people from business and industry fields, universities, municipalities, and civil society). Different countries have their own national strategies. For example, Finland published its first AI strategy in 2017 (Ministry of Economic Affairs and Employment in Finland, MEAE 2017) and has provided updates (MEAE 2019). The main goal of the guidelines is to benefit from the opportunities brought by AI in all areas of society but in such a way that ethical aspects are considered and possible risks avoided. The Organisation for Economic Co-operation and Development (OECD) *Recommendation of the Council on Artificial Intelligence* (OECD 2019) has listed more than 70 documents published in the last 3 years which make recommendations about the ethics principles for AI (Spielkamp et al. 2019; Winfield 2019).

It is noteworthy that most of the guidelines developed by companies and other organisations focus on what ethical challenges exist, rather than what actions should be taken to achieve the ethical goals in practice (Cath 2018; Morley et al. 2020). It has been argued that the developers are often aware of the ethical issues, but companies do not provide appropriate tools or support to suitably tackle these issues (Abdul et al. 2018). Ethical guidelines for education as a context of AI application are mainly lacking (Holmes et al. 2021), although the need has been recognised decades ago (Aiken and Epstein 2000). Nonetheless, educational issues are included in general policy-level guidelines (e.g. AI HLEG 2019a). Jobin et al. (2019) analysed 84 regulation documents or guidelines for ethical use of AI, and according to their review, the most important principles are transparency (including explainability and understandability), justice and fairness, non-maleficence, responsibility, and privacy. In addition, Hagendorff (2020) has presented ethical criteria such as accountability, explainability, discrimination-aware data mining, tools for bias mitigation, and fairness in machine learning. Moreover, AI actions should also be predictable and the systems that are based on AI should be robust against manipulation. Clear human accountability for AI actions must also be ensured (Bostrom and Yudkowsky 2011).

According to a literature review by Morley et al. (2020), the five main principles are beneficence, non-maleficence, autonomy, justice, and explicability, which are not only complementary but also partly overlapping. Morley et al. (2020) have combined this typology from the EU's report that lays grounds for trustworthiness (AI HLEG 2019a). The five principles can be summarised as follows:

- *Beneficence* means that the AI-based system is useful and reliable and supports diversity, human well-being, and development. Product development should be justified in alignment with beneficence and not created solely for the sake of the product but for benefiting the user.
- *Non-maleficence* includes many aspects that are related to human and data safety. It is important to be prepared in advance for the possible security threats. Data

security, accuracy, reliability, reproducibility, quality, and integrity must each be guaranteed at all stages of the product's life cycle.

- Autonomy means the freedom to make decisions and choices regarding AI-based systems. Tools that support individual autonomy should be designed and implemented.
- Justice requires that the AI systems operate in a fair manner without obstructing democracy or harming society. The negative effects of AI systems should be minimised. All stages of product development processes should be made more transparent, and they should be able to be evaluated and documented.
- Explicability means that AI systems should be able to be understood and their operations explained and interpreted. This does not mean that everything should be explained, as it is impossible. The level at which AI is explained depends on need and can range from a very simple explanation to a more complex one. For example, basic knowledge could be taught and assessed in schools in ways appropriate to different age levels. This would ensure sufficient civic skills to make ethically sustainable decisions, for example, regarding one's own security. Accountability and responsibility should be clear, transparent, and traceable.

The typology introduced by Morley et al. (2020) shows that many of the ethical principles are very interrelated. Explicability can be seen as both an independent and a unifying factor. In many cases, it is unclear what needs to be explained concerning AI and its applications and how the decision is made (Coeckelbergh 2020) and who makes the decisions (Floridi et al. 2018). Additionally, it is not always clear who should take responsibility if something goes wrong or if AI is to blame in those occasions. In the next section, representatives of EdTech companies and schools will reflect on what their major concerns from an ethical viewpoint are when AI is applied in educational settings.

4 Case I: Finnish EdTech Companies' Views on Ethical Challenges

Seven EdTech company representatives who work in Finland were interviewed in the qualitative study of Kousa and Niemi (2021). The aim was to look for new ideas and solutions on how AI could be utilised in an ethically sustainable way in education. Companies in this study provide AI- based EdTech products and services such as well-being surveys and solutions for schools, tutoring services using VR and AR technology, ethical and safe data management solutions, and game- and simulation-based applications in oil operator training. All companies have extensive international business and more than 10 years of experience in the EdTech field. According to the findings, EdTech companies have faced ethical challenges in their work.

First of all, companies struggle with regulations and guidelines which have been found difficult to understand and implement. Therefore, making their own

guidelines is mostly preferred. The situation is even more complicated in the international marketplace for educational technologies, since other countries are likely to have different cultures, guidelines, and understandings for what is meant by ethical AI in the first place. Additionally, conducting business with schools is challenging as schools' resources, opportunities, and willingness to use AI-assisted solutions vary widely. Negative attitudes or even unrealistic expectations of AI were also seen as problematic. The situation is contradictory when, on the one hand, information is freely provided, for example, on social media, but, on the other hand, there are many kinds of fears. For example, AI solutions are not necessarily trusted in the workplaces or schools, or workers might be afraid that machines will replace them in the future. It was also argued that the bad reputation and negative attitudes of AI is caused by the critical tone with which large companies such as Microsoft, Google, Apple, Facebook, and Amazon have been talked about in the media.

When EdTech companies were asked how they could increase ethical sustainability in their AI solutions, the following issues were raised:

1. Companies argued that there is a serious lack of civic knowledge about what AI is and what it can and cannot do. Therefore, more public education about AI in education is needed. For example, AI is often misunderstood to mean just learning coding in schools. Therefore, AI as well as related ethical issues should be taught at all grade levels. Teachers also need more education on the topic.
2. Schools and workplaces should have equal opportunities to choose AI-based teaching materials and methods that are accessible and understandable. However, customised versions of one-size-fits-all services that are increasingly needed are expensive and often impossible to implement. Therefore, the best solutions could be based on collaboration between the teacher and the AI application, such as an AI tutor.
3. Responsibility issues should be defined and more transparent. It should be clearer to the user when the company is responsible and what responsibilities belong to the user. However, it is typically unclear when a machine or human is responsible when AI is applied in education.
4. One of the companies' main concerns was how to make safe and ethically sustainable products for schools. Problems were seen in data collection, transfer, storage, and modification. Prevention of harm is one of the most important solutions to make more ethically sustainable solutions. That means constant risk analysis and ethical checklists.
5. Universities, companies, and schools should share more of their knowledge and best practices to each other and to the public for a common good about AI in education.
6. Public events about the possibilities, risks, and threats of AI should be held regularly to avoid development of false assumptions about AI and formation of only negative attitudes.

When EdTech companies were asked about the need for support, several issues surfaced. Companies need more understanding about legislation, ethical risks, algorithms, and responsibility issues. They hope that there would be multi-professional

partners such as legal experts, universities, schools, other companies, or decision-makers who could be asked for support and advice on difficult ethical issues. They also wanted to share responsibilities between different stakeholders.

5 Case II: Finnish School's Ethical Challenges and Practical Viewpoints on Explicability

Twenty school principals and/or teachers who work as digital tutors in Finnish schools participated in a qualitative interview study in 2021. The participants were asked about their views on AI, digital applications, and ethical challenges.

As for what constitutes the main challenges related to AI in education, many respondents felt that teachers do not know enough about AI or related applications. According to interviewees, there are usually only a few more dedicated teachers in schools who act as digital educators/tutors. One of the school principals stated that teachers are not motivated to adopt AI tools if there is no guarantee that they will be useful in teaching. In smaller schools, the acquisition and responsibility for digital equipment was generally the responsibility of the principal. AI was seen as a good tool for easy routine tasks and for providing differentiated instructions when needed. However, all teachers did not see AI or digital applications alone as sufficient to guarantee better teaching or learning. One teacher described the scenario as follows: 'So the AI would say to the teacher that Matt is a bit stressed now so you should leave him alone (laughter)? I have to say that I can't imagine what kind of help AI could provide that a teacher cannot. Even though it is AI, someone has coded it. There should also be some kind of control that AI gives the right information before we start doing things based on it'. When asked what kind of additional information teachers would need about AI, one replied: 'We should find out what AI means in practice. If we have an application that collects information about stress, then we need to know well enough about its operating principles and purposes. To see the big picture. And what to think about AI in education'. According to teachers, AI-based applications should be developed in collaboration with schools, companies, and researchers and should be tested long enough before use. One of the future scenarios which teachers are afraid of is that when the use of AI-based solutions increases, their control in the classrooms will diminish. They are concerned that companies are starting to define more and more about what is taught and how. This in turn might reduce objectivity as one of the teachers explained: 'I hope that we would get better AI tools for teaching. This means that our city, which decides what tools are allowed, has to reduce strict restrictions, and make more new contracts with different software houses. Then there is a fear that it will go so that there will be those lobbyists of big companies such as Microsoft, which will forge them. I think our city has a fear that schools would be in breach of EU regulations if they were allowed to decide for themselves which AI tool to use.' In another example, the teacher expressed the concern: 'If AI begins to define what individual students do

in the lesson based on their personal learning profiles, the situation is not controlled by the student or a teacher or the parents, but by some other parties.' Furthermore, teachers did not believe that even the smartest system could replace teachers or make equally good predictions about how to work with diverse students or make decisions for the benefit of students. 'When thinking about an entire school day, it will always be influenced by a terrible number of elements that are related to only one situation. Predicting them and drawing more long-term conclusions would seem to be quite difficult, at least for the time being. For example, we know that in the fall, when it rains and is dark, disruptive behaviour easily increases. In this case, classroom lightning and human factors such as teacher's situational awareness are of great importance. If interpretations, conclusions, and measures come through AI, we will go to the so-called schematic side. That's when we're lost the human side of teaching.'

School representatives also felt that they have unequal opportunities to use AI-based solutions. The situation differs enormously even within a city. Some of the interviewees argued that there are schools that do not even have proper Internet connections and there are teachers and parents who are against digital education since they are afraid of, for example, issues related to privacy or even health.

Information security was an important issue in interview discussion, but there were differences in teachers' opinions on this topic. Some were not worried about sharing information, and others were very precise and also knew about the importance of privacy issues. However, security was seen as a challenge that companies and/or the city needs to address and cannot be an individual teacher's responsibility. New applications and unknown, especially foreign, companies were seen as less trustworthy. Indeed, many believed that larger companies had taken better care of information security. To improve safety, it was proposed that data taken from students should be stored only for a short time and then safely disposed of. Other options they proposed were that students' data should be anonymised and stored in an encrypted form in a secure location so that no one could recognise the student from the data. When asked about the future scenarios, one of the interviewees summarised: 'After all, the school is not out of the community. And AI comes into society on a global scale, whatever was said or done in schools. However, the school should not be the first place to use AI for the industry or business purposes, but vice versa. Schools need to keep up with the development of technology on their own terms. It is challenging because the changes are happening at an increasingly hectic pace. In schools, we need to remember that we are dealing with children or young adults. It seems like we have forgotten the stages of Piaget's cognitive development and so on. It seems that sometimes children are being expected too much these days'.

In order to facilitate the situation in schools where digital skills are becoming more important and a wide range of programs are used and provided by EdTech companies, the interviewees stated that:

1. It has to be explained what kind of data the system collects and what it is used for. In addition, teachers want to know where the data is stored and who owns

    it. This increases the credibility of the system in question and even influences its purchasing decision.

2. Teachers prefer easy-to-use and effective teaching aids. Some of the teachers even argued that privacy or explicability are not as important as usability.

3. It is difficult to teach with digital applications without understanding how they work. This was noticed especially during the COVID-19 pandemic in distance teaching. Digital applications need to add value to teaching. Teachers have found that there are too many applications where traditional teaching has been digitised without major changes. These applications can make learning and understanding even more difficult.

4. Many schools have IT support which is usually one person who works at school or in a larger area that includes many schools. According to teachers, sufficient help is mostly difficult to get. One of the problems is that IT supporters rarely know about the pedagogical aspects of the system and therefore cannot explain how it works or what kind of features there are that could help teaching or learning. Therefore, EdTech companies' support was seen as essential.

5. Although attitudes towards AI are generally positive, most of the teachers have noticed that they have insufficient basic knowledge about it. For example, many teachers did not see any differences between coding and understanding basic things about AI to achieve adequate civic skills for the future.

6. Teachers interviewed have no resources or desire to learn about AI or to learn numerous new AI systems, although ICT is expected to be used at all stages of education in Finland. Most of the respondents hoped that AIED would be used in teaching only by voluntary teachers who were interested in it.

7. Decision-makers at schools prefer digital solutions that are as accessible and understandable to as many teachers as possible.

8. Teachers do not want to be responsible for the privacy issues or functionality of the digital/AI-based solutions but assume that either the companies or the municipality are responsible. According to teachers, liability issues concerning EdTech products should be regulated by law.

9. Teachers hoped that someone, such as a municipality, would review all EdTech companies and their products and services and ensure that they are ethical and safe to use.

6   Discussion

Ethical issues are strongly present in the daily lives of both schools and companies. These two cases represent a small sample of the situation in Finland, where the technological skills and know-how are at an internationally high level. However, more information is needed on the ethical issues involved and how the gap between businesses and schools could be reduced, inter alia, to improve trustworthiness. EdTech companies' and schools' challenges are discussed in the light of five ethical principles (beneficence, non-maleficence, autonomy, justice, and explicability) by Morley et al. (2020).

6.1 Beneficence

In Morley et al's typology, beneficence means that AI brings something positive to users and community and that AI is not a purpose in itself. Teachers strongly emphasised that the use of AI programs would not be an absolute value but would be based on a genuine need, for example, for differentiated instruction or assisting with routine tasks. Since teachers have a constant shortage of time and money, beneficence is an extremely important factor in choosing the right tools for teaching and learning. Companies also see the importance of providing accessible systems that take diversity into account, but they are worried that providing customised versions of one-size-fits-all solutions is challenging. Morley et al. (2020) see that justification belongs to beneficence. The purpose for building the system must be clear and linked to a clear benefit—systems should not be built simply for the sake of AI application or profit only.

6.2 Non-maleficence and Justice

Non-maleficence and justice are very much interrelated in the conceptualization of Morley et al. (2020): it means that AI systems should be protected against vulnerabilities that can allow them to be exploited by adversaries. AI systems should have safeguards that enable a fallback plan in case of problems. AI systems should guarantee privacy and data protection throughout a system's entire life cycle. Justice requires minimising and responding to potential negative impacts of AI systems. Companies on this study want to avoid ethical risks and emphasise that they do not intentionally make AI solutions that would be harmful to an individual or society. However, they need proper guidance, information, and legislation to support their product development processes. On the other hand, schools also need proper guidance on how to safely use digital/AI-based solutions. Recent research of Felderer and Ramler (2021) brings up the importance of quality assurance of AI-based systems. It has been recognised by AI solution developers that the models of machine learning or deep learning are not transparent, intuitive, or understandable. In Europe, General Data Protection Regulation (GDPR) has been developed to understand data management processes and civil rights, for example, how to protect users' personal data (EC 2018). However, identifying the factors that make AI non-maleficent requires considerable understanding of the entire system, from both developers and users. According to EC (2021) people should have basic digital skills and knowledge of AI and the ability to access and use the solutions in their daily lives.

   According to one expert group (AI HLEG 2019a), accountability includes 'auditability, minimization and reporting of negative impact, trade-offs and redress' (p. 14). It is related to fairness and responsibility which are extremely necessary in every step of the production development process, both before and after. In this

study, companies emphasised that systems should be preventive of and minimise the risks. Companies complained about the difficulty of legislation and preferred their own guidelines and checklists. Hagendorff (2020) argues that ethical guidelines might not have a sufficient impact on companies' decision-making. They can be interpreted in many different ways because concepts are not clear. It is also easy to slip up on adherence to ethical principles, since there will be no consequences, surfacing policy concerns.

6.3 Autonomy

Autonomy means human agency and human oversight in a typology of Morley et al. (2020). This means that even though machines can intelligently analyse data and make conclusions, human beings are still responsible for the system and its consequences. Teachers in this study admitted that they do not want to be responsible for the privacy issues or functionality of the digital/AI-based solutions. They do not have the capacity to accomplish that. They assumed that either companies or the municipality should be responsible. The situation was twofold in these cases. Companies, on the other hand, understood their responsibilities but also wanted to share them among different stakeholders.

6.4 Explicability

Morley et al. (2020) set us an aim that AI systems should be built in such a way that they are understandable to users. Companies in this study needed more education and knowledge sharing to increase public trustworthiness in AI and its applications. Schools also needed information on both AI and its applications. Coeckelbergh (2020) points out that without explainability and transparency, responsible use of AI technology is problematic. To act in an ethically responsible way means knowing what is being done and being able to explain the system's actions and decisions in a way that others can understand. In addition, it is important to know to whom one is responsible for the creation of AI systems. The issue is complex, because people's need for explanations varies. Most people don't necessarily know that AI is involved in their applications in the first place or what AI does in that application. Even the best software developer may not know all the codes or know how to explain them (Coeckelbergh 2020). It can be concluded that explainability is a very human, content-, and context-dependent issue and, therefore, while extremely complicated, necessary.

7   Conclusions

This chapter has discussed the ethical challenges of EdTech companies and schools. Although EdTech companies and schools share some challenges, it can be said that the gap between companies and schools is in danger of widening as technological development advances. This observation also applies to other parties in the society, including researchers, decision-makers, and legislators. First, in the absence of suf-ficient legislation in the AIED field (Aiken and Epstein 2000; Holmes et al. 2021), ways should be urgently found for how to develop globally consistent regulations and guidelines, which include practical examples in a sufficiently understandable way to meet educational needs. This topic requires further research and consultation with both parties, as well as legitimate solutions based on consensus. Secondly, it must be recognised that explicability is a broad concept with many levels and needs (e.g. decision-makers, developers, and users) including what needs to be explained and how. In addition to understanding the technical details of individual applications and 'black boxes,' more knowledge is needed concerning how to explain AI in general and in the context of everyday life implementations. As stated earlier, it is not necessary to explain everything (Coeckelbergh 2020), but it is, for example, necessary to obtain the necessary civic knowledge and skills to participate in society. That could mean, for example, specifying what added value AI brings to the application used by the teacher or how. In conclusion, a huge amount of work has been done by researchers, companies, policymakers, and schools to increase a common understanding of AI. However, we are still on our journey to a more ethically sustainable future.

Disclosure Statement   No potential conflict of interest was reported by the authors.

References

Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018). Trends and trajectories for explainable, accountable, and intelligible systems: An HCI research agenda. In Proceedings of the 2018 CHI conference on human factors in computing systems—CHI'18 (pp. 1–18). https://doi.org/10.1145/3173574.3174156.

Aiken, R., & Epstein, R. (2000). Ethical Guidelines for AI in Education: Starting a Conversation. *International Journal of Artificial Intelligence in Education,* 11, 163–176.

Bostrom, N., & Yudkowsky, E. (2011). The ethics of artificial intelligence. In K. Frankish & W. Ramsey (Eds.), *The Cambridge Handbook of Artificial Intelligence* (pp. 316–334). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139046855.020.

Cath, C. (2018). Governing artificial intelligence: ethical, legal and technical opportunities and challenges. Phil. Trans. R. Soc. A376: 20180080. http://dx.doi.org/10.1098/rsta.2018.0080.

Coeckelbergh, M. (2020). Artificial intelligence, responsibility attribution, and a relational jus-tification of explainability. *Science and Engineering Ethics*, 26, 2051–2068. https://doi.org/10.1007/s11948-019-00146-8.

Dignum, V. (2018). Ethics and artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20, 1–3. https://doi.org/10.1007/s10676-018-9450-z.

EC (European Commission) (2018). European Commission for the Efficiency of Justice (CEPEJ). European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment. Strasbourg, CEPEJ. Retrieved October 12, 2021 from https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.

EC (European Commission) (2021). Digital skills and jobs. Retrieved October 12, 2021, from https://digital-strategy.ec.europa.eu/policies/digital-skills-and-jobs

European Commission AI HLEG (High-Level Expert Group on Artificial Intelligence) (2019a). Ethics guidelines for trustworthy AI. https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top. Accessed April 8, 2019.

European Commission AI HLEG (High-Level Expert Group on Artificial Intelligence) (2019b). Policy and investment recommendations for trustworthy AI. Accessed June 26, 2019, https://futurium.ec.europa.eu/en/european-ai-alliance/open-library/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

European Commission AI HLEG (High-Level Expert Group on Artificial Intelligence) (2020a). The Assessment List for Trustworthy Artificial Intelligence (ALTAI). https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342. Accessed July 17, 2020.

European Commission AI HLEG (High-Level Expert Group on Artificial Intelligence). (2020b). Sectoral Considerations on the Policy and Investment Recommendations for Trustworthy Artificial Intelligence. https://futurium.ec.europa.eu/en/european-ai-alliance/document/ai-hleg-sectoral-considerations-policy-and-investment-recommendations-trustworthy-ai. Accessed July 23, 2020.

Felderer, M., & Ramler, R. (2021). Quality Assurance for AI-Based Systems: Overview and Challenges (Introduction to Interactive Session). In D. Winkler, S. Biffl, D. Mendez, M. Wimmer, & J. Bergsmann (Eds.), *Software Quality: Future Perspectives on Software Engineering Quality. SWQD 2021.* Lecture Notes in Business Information Processing, vol. 404. Springer. https://doi.org/10.1007/978-3-030-65854-0_3.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V. Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28, 689–707. https://doi.org/10.1007/s11023-018-9482-5.

Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., Kieseberg, P., & Holzinger, A. (2018). Explainable AI: The new 42? In A. Holzinger, P. Kieseberg, A. Tjoa, E. Weippl (Eds.), *Machine learning and knowledge extraction.* Lecture Notes in Computer Science, vol. 11015 (pp. 295–303). Springer. https://doi.org/10.1007/978-3-319-99740-7_21.

Grover, S., & Pea, R. (2018). Computational thinking: A competency whose time has come. In S. Sentance, E. Barendsen, & C. Schulte (Eds.) *Computer science education: Perspectives on teaching and learning* (pp. 19–38). London: Bloomsbury Academic.

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8.

HAI (Human-Centred Artificial Intelligence). (2020). Human Impact Research Mission. Stanford University. Retrieved October 12, 2021, from https://hai.stanford.edu/research/human-impact.

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O., Rodrigo, M., Cukurova, M., Bittencourt, I., & Koedinger, K. (2021). Ethics of AI in Education: Towards a Community-Wide Framework. https://doi.org/10.1007/s40593-021-00239-1.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kousa, P., & Niemi, H. (2021). AI ethics and learning – educational technology companies' challenges and solutions. *Interactive Learning Environments* (Submitted in final review).

Littman, M., Ajunwa, I., Berger, G., Boutilier, C., Currie, M., Doshi-Velez, F., Hadfield, G., Horowitz, M., Isbell, C., Kitano, H., Levy, K., Lyons, T., Mitchell, M., Shah, J., Sloman, S., Vallor, S., & Walsh. T. (2021). *Gathering Strength, Gathering Storms: The One Hundred Year*

Study on Artificial Intelligence (AI100) 2021 Study Panel Report. Stanford University, Stanford: CA. Retrieved September 16, 2021, from http://ai100.stanford.edu/2021-report.

MEAE (Ministry of Economic Affairs and Employment of Finland) (2019) Leading the way into the age of artificial intelligence: Final report of Finland's Artificial Intelligence Programme 2019. Retrieved from October, 12, 2021, from http://urn.fi/URN:ISBN:978-952-327-437-2

MEAE (Ministry of Economic Affairs and Employment, Finland) (2017). Turning Finland into a leading country in the application of artificial intelligence: Objective and recommendations for measures. Retrieved from October, 12, 2021, from https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1amp;isAllowed=y

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). 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, 2141–2168. https://doi.org/10.1007/s11948-019-00165-5.

Niemi, H., & Kousa, P. (2020). A case study of students' and teachers' perceptions in a Finnish high school during the COVID pandemic. *International Journal of Technology in Education and Science (IJTES)*, 4(4), 352–369. https://doi.org/10.46328/ijtes.v4i4.167.

Nnaji, J. (2019). Ethical dimensions of the increasing usage of new technologies in virtual education. In A. Blackburn, Chen, L, & Pfeffer, R. (Eds) *Emerging Trends in Cyber Ethics and Education* (pp. 1–21). USA: IGI Global.

OECD (2019). Recommendation of the Council on Artificial Intelligence. Accessed May 22, 2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

Spielkamp, M., Matzat, L., Penner, K., Thummler, M., Thiel, V., Gießler, S., & Eisenhauer, A. (2019). Algorithm Watch 2019: The AI Ethics Guidelines Global Inventory. Accessed April 9, 2019, https://algorithmwatch.org/en/project/ai-ethics-guidelines-global-inventory/

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M., & Teller, A. (2016). Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA. http://ai100.stanford.edu/2016-report.

Torresen, J. (2018). A Review of Future and Ethical Perspectives of Robotics and AI. *Front. Robot. AI*, 4(75). https://doi.org/10.3389/frobt.2017.00075.

Turing, A. M. (1950/2009). Computing machinery and intelligence. *Mind*, LIX (236), 433–460. https://doi.org/10.1093/mind/LIX.236.433.

UNESCO (2019). Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development. Retrieved from January 5, 2022, from https://unesdoc.unesco.org/ark:/48223/pf0000366994.

Winfield, A. (2019). An updated round up of ethical principles of robotics and AI. Accessed April 18, 2019, http://alanwinfield.blogspot.com/2019/04/an-updated-round-up-of-ethical.html.