PREPARATORY STUDY FOR THE DEVELOPMENT OF A LEGAL INSTRUMENT ON REGULATING THE USE OF ARTIFICIAL INTELLIGENCE SYSTEMS IN EDUCATION

REVISED DRAFT

Directorate General of Democracy and Human Dignity
Directorate for Democracy

Education Department
Digital Transformation Unit

CONTENTS

AKNOWLEDGEMENTS

1   EXECUTIVE SUMMARY

The preparatory study highlights the increasing use of Artificial Intelligence (AI) technologies in educational systems and their potential to influence educational practices. The Council of Europe's initiative aims to develop a legal instrument to regulate AI systems in order to safeguard human rights, democracy and the rule of law in educational settings.

The use of AI systems in education, accelerated by the arrival of generative AI tools such as ChatGPT, encompasses a wide range of technologies, from adaptive tutoring systems to plagiarism detection and e-proctoring, with implications for students, teachers and institutional frameworks. Despite their growing presence, there is a worrying lack of robust, independent evidence about the effectiveness, safety and overall impact of these technologies in educational contexts. In addition, many AI tools, particularly those developed for commercial purposes, lack transparency, contributing to unevaluated, potentially unsafe applications in education and thus compromising human dignity and student rights.

Children and education are unique cases. Children, due to their stage of development, have specific needs and rights that distinguish them from adults. Similarly, education, as a formative element of human development, has unique challenges and responsibilities. Unfortunately, current AI policies and legislation generally overlook these specific needs.

AI systems, including those not specifically designed for educational settings, can exacerbate existing biases, violate privacy, embed naïve pedagogies, and fail to recognise the unique developmental stages of children, with potentially lifelong consequences. This misuse highlights the need for strong, child-focused regulations in the procurement and use of educational technology.

Furthermore, the use of AI systems risks widening the digital divide, disproportionately affecting marginalised groups by limiting access to educational resources and opportunities. Despite these challenges, current regulations are insufficient to address the specific needs and rights of children and educational institutions.

The document argues for new, comprehensive regulations that focus on protecting learners and educators and maintaining the integrity of the education system amid the growing influence of AI. It emphasises the need for an ecosystem approach to regulation, involving multiple stakeholders and balancing innovation with protection.

The study suggests several regulatory needs, including transparency in the development and use of AI systems, accountability frameworks for AI tools in education, and mechanisms to mitigate potential harm from inappropriate content or biased algorithms. It also calls for international cooperation to harmonise regulation of AI in education, promote cross-border collaboration, and ensure consistent application of human rights principles.

In summary, while AI potentially offers opportunities to improve educational outcomes, its application requires careful regulation to prevent harm, protect human rights and ensure equitable access to educational benefits. The proposed legal instrument aims to address these challenges by providing a structured approach to the integration of AI in educational contexts, while safeguarding the interests of all stakeholders, in particular children.

1.1   Overview of the main chapters and sections

Introduction

The initial section introduces the Council of Europe's proposition for a legal framework aimed at regulating AI systems within educational environments. It touches on the potential benefits and risks associated with AI in education, the lack of robust evidence supporting the safe and effective use of AI technologies, and the unique requirements and rights of children within educational settings. The section emphasizes the need for comprehensive legislation that

addresses these unique challenges and ensures the protection and promotion of human rights, democracy, and the rule of law.

Background

This chapter elaborates on the increasing incorporation of AI in educational systems. It highlights the diversity of AI applications in education and the emerging issues such as lack of transparency, accountability, and evidence regarding their efficacy and safety. The section dives deeper into:

AI in Education: Overview of AI technologies in educational settings, emphasizing the absence of substantial evidence for their effectiveness and potential risks.

Unique considerations for children and education: Discusses the particular needs and rights of children and how education serves as a pivotal element of human development, arguing that current AI strategies overlook these specifics.

Specific areas of concern: Focuses on human rights, consent, procurement, commercial influence, academic integrity, adaptive learning, and areas of high risk specifically related to education.

Regulatory aspects: Covers various dimensions of regulation, including government intervention, principles for the use of AI systems in education, self-regulation, standards, and the need for widespread ecosystem support.

Questions for discussion:

This section presents a series of questions aimed at prompting reflection and dialogue regarding the regulation of AI systems in education.

These questions explore prioritisation areas within human rights, democracy, and the rule of law, ethical norms, mechanisms for ensuring AI's positive impact in classrooms, acceptable levels of risk, and specific use cases needing special attention.

It further discusses the implications for stakeholder relationships, the digital divide, harmonization of regulations across member states, and the roles of various stakeholders in implementing the proposed legal framework.

2 INTRODUCTION

The Council of Europe Standing Conference of Ministers of Education (26th Session, September 2023) welcomed the proposal for developing a legal instrument to regulate the use of Artificial Intelligence systems in education to promote and to protect human rights, democracy, and the rule of law. In this document, we set out the case for the proposed legal instrument.

Artificial Intelligence (AI)-enabled technologies have been widely proposed as a means by which children's human rights to education might be achieved. However, without appropriate safeguards, such technologies might instead threaten such rights.

1. **For more than a decade but without much public awareness, AI-enabled technologies have increasingly been used in education. This has been accelerated by the arrival of ChatGPT and other generative AI.**

As discussed in the Council of Europe report, "*Artificial Intelligence and Education. A Critical View Through the Lens of Human Rights, Democracy, and the Rule of Law"*, AI-enabled technologies (AIED) have been researched in educational settings for more than forty years. These technologies include **student-focused AIED** (from adaptive tutoring systems to support chatbots), **teacher-focused AIED** (from plagiarism detection to resources curation), and **institution-focused AIED** (from admissions to e-proctoring).[1]

Over the past decade, many of these tools have become commercial products that are increasingly being sold into educational institutions. Although these products are typically 'black boxes' (neither their decision-making processes nor their decisions are open to inspection), they are frequently granted influence or control over the provision, delivery, and access to education. Other AI-enabled technologies that have not been designed for use in education are also increasingly being used in educational settings (e.g., Microsoft Office, Google Docs, and generative AI). Without mechanisms to adequately evaluate all these different types of technologies, and because they exist outside of any democratic control or process, there is no way to assess and ensure the claims made by AI providers, their safe and private use of data, their ability to uphold human rights, or their impact on the current and future lives of learners.

The widening adoption of AIED in educational contexts has been accelerated by the dramatic arrival of generative AI (following the launch of ChatGPT), which has led to a deluge of novel education-targeted AI-enabled applications (from lesson planning to idea generation). While some of these tools assign students' individual accounts to the generative AI tool directly, others build on top of other products. This means that learners using these applications are unknowingly both indirect users of and used by the back-end AI-enabled system. **However, although much public discussion today does focus on generative AI, there are at least twenty-five different categories of AIED, of which generative AI is only one.[2]**

2. **There is no robust independent evidence at scale for the efficacy, safety or positive impact of AI systems being used in education.**

While the AI in Education research community have undertaken thousands of studies, **the vast majority of those studies assess only efficacy** (e.g., Does this particular tool, in this particular context, improve student test scores?). **Rarely do they consider the broader safety of the tool** (e.g., What is the impact of such tools on the mental health, privacy, or agency of the students and teachers?); and **they almost never consider the impact of the tool on the educational ecosystem** (e.g., What is the impact of the tool's deployment on the

---

1. Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education: Research, Development and Policies, 57(4), 542–570. https://doi.org/10.1111/ejed.12533

2. Holmes, W. (2023). The Unintended Consequences of Artificial Intelligence and Education. Education International Research. https://www.ei-ie.org/en/item/28115:the-unintended-consequences-of-artificial-intelligence-and-education

classroom, on relationships between teachers and students, and on the professional skills of teachers?). In addition, claims made about AI systems designed for education are often unevidenced. Examples include AI systems that claim to detect autism to 96% accuracy[3] or to predict pupils' mental health and wellbeing (involving issues such as self-harm, drug abuse and eating disorders), and to propose interventions.[4] As 'education' tools, none of these systems are audited or have oversight.

In fact, there are currently no comprehensive or widely accepted mechanisms or protocols to measure and assess the efficacy, safety or wider impact of AI systems being used in education.[5] Very few of the existing studies have been conducted independently, undermining the integrity of the studies, and very few have been undertaken at scale, undermining their credibility. There is, however, growing evidence of the harm that AI systems might cause in educational settings. For example, many AI-enabled educational technologies, by default, carry out some degree of social scoring (comparing behaviours, characteristics and outcomes of learners and educators). Meanwhile, some AI-enabled online examination proctoring often unfairly prevents students taking their exams. They can discriminate against some students due to the colour of their skin or a disability and can exacerbate mental health problems.[6] In short, the imposed adoption of such products can interfere with the students' human dignity.

3. Children and education are special cases.

Children, like all humans, are individuals with unique needs, vulnerabilities, and potential. However, children are also a special case because of their physical, emotional, cognitive, and spiritual immaturity. For these and other reasons, children have both human rights (ECHR) and child rights (UNCRC). Violations of children's rights (by exposure to harm, child labour, environmental hazards, unsafe products, unsafe practices, and commercial exploitation) may have lifelong, irreversible, and even transgenerational consequences.

Most national AI strategies and major ethical guidelines make only cursory mention of children and their specific needs. The new ways that children and young people interact with AI-supported technology have implications to physiology and psychology yet to be fully understood (UNICEF, 2018).[7] How does interaction with voice-operated tools shape how children speak, communicate and socialise? Do children develop relationships[8] with AI embodied in humanoid robot reading assistants?[9] How might these interactions influence their expectations of gender-based norms?

Education is a special case because of its role in actively shaping and supporting human development. More specifically, the aims of education are to enable a child to develop their full potential (UNCRC, Article 29), and for the development of informed, responsible citizens who actively participate in democratic processes and address societal challenges. The application of AI can interfere with these aims. The education environment also holds a special status because children are required to be in suitable education that means being 'in school' is not a consensual choice in the same way as it might be for adult learners across member states, and in line with SDG 4.

3. Hill, J. (2020) Schools Week https://schoolsweek.co.uk/ceos-mission-to-make-ai-a-force-for-good-in-classrooms/

4. Sky News (2019) https://news.sky.com/story/artificial-intelligence-being-used-in-schools-to-detect-self-harm-and-bullying-11815865

5. USA Office of Educational Technology. 'Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations' (2023) https://www2.ed.gov/documents/ai-report/ai-report.pdf

6. Why is online proctoring under fire? (2023) https://www.onlineeducation.com/features/online-exam-proctoring

7. UNICEF (2018) Children and AI https://www.unicef.org/globalinsight/featured-projects/ai-children | Where are the opportunities and risks? https://www.unicef.org/innovation/sites/unicef.org.innovation/files/2018-11/Children%20and%20AI_Short%20Verson%20%283%29.pdf

8. Kahn, P. H., Jr., Gary, H. E., & Shen, S. (2013). Children's social relationships with current and near-future robots. *Child Development Perspectives, 7*(1), 32–37. https://doi.org/10.1111/cdep.12011

9. VU Amsterdam (2023) Fighting the reading crisis with a reading robot https://vu.nl/en/news/2023/fighting-the-reading-crisis-with-a-reading-robot

4. **The possible impact of AI systems on children and teachers includes generic issues (focused on data, privacy, and biases). Existing legislation sets out to address most of these generic AI issues but does not specifically address the particular impact of such issues on children or education settings.**

It is well-known that all AI systems are affected by a range of critical issues – from profiling to privacy, biases to misuse of data, and environmental impact to discrimination by race or gender. Accordingly, in most countries there is already legislation (such as GDPR in Europe) that sets out to address broader, societal issues. However, this legislation rarely acknowledges the special case of children (the fact that childhood is a unique period of physical, emotional, cognitive, and spiritual development) or the special case of education.

In particular, AI-enabled technologies are rarely trained on children's data (partly due to the need to protect children's biometric data and identity). Instead, they are usually trained on adult data, such that the derived models and weightings may be inappropriate for children[10] or based on adult or household rather than individual profiling.[11] In addition, **no child-appropriate assessments have been undertaken into the environmental, mental health, human rights impact of AI-enabled technologies designed for education**; and there is no evidence of consistent and appropriate standards for auditing harms in education procurement processes.

In an important sense, education technologies raise parallel concerns to those in medicine. Just as drugs and medical treatments are mostly designed to impact on the human body, so AI-enabled education technologies are designed to impact on the human mind and especially on a child's developing cognition. If they do not impact, there is little reason for using them.[12] However, key practices found in health are still undefined in education. For example, before being available for use, medicines undergo a rigorous stepped safety trial process, which evaluates both efficacy and safety. Inevitably, that process is not perfect; bad drugs do sometimes get through the system and there is also off-label use. However, robust testing gives doctors confidence that their prescriptions will address a particular set of symptoms and will do so safely. For AI-enabled education technologies, almost no such testing takes place, which means that teachers must rely on anecdotal evidence or marketing claims.

5. **The possible impact of AI-enabled technologies on children and teachers also includes education-specific issues (such as choice of pedagogy, teacher disempowerment, and student agency).**

In addition to the generic issues mentioned above, all AI-enabled educational technologies also raise but rarely address education-specific issues. These include the quality of education, choice of pedagogy (whether the embedded pedagogy is properly evidenced), teacher disempowerment (whether the technology takes control away from teachers), the appropriateness of approaches to assessment, learner's over-reliance on technology (undermining, for example, writing and critical thinking skills), access and inclusion, social interaction, and student agency. While all these education-specific issues predate AI, the use of AI-enabled technologies in education can exacerbate each of them.

The appropriate choice of pedagogy, for example, can enable a child to reach their full potential (as specified in UNCRC Article 29). However, there is currently a lack of oversight and transparent regulation to ensure that AI-enabled educational tools do not just automate poor educational practice, but rather ensure a child's intellectual, emotional and social development.

---

10. Shahnawazuddin, s. et al (2021) robust children's speech recognition in zero resource condition https://www.sciencedirect.com/science/article/abs/pii/s0003682x2100476x

11. Barassi, v. (2020) the human error in ai and question about children's rights https://childdatacitizen.com/wp-content/uploads/2020/06/the-human-error-in-ai-and-children-rights_prof.-barassi_response-to-ai-white-paper-.pdf

12. *"The process of testing pharmaceutical products must comply with strict regulations before entering the consumer market. It seems that similar rules do not hold for computer technology products. Is it because we believe that computer technology does not affect our bodies or psyche?"* Martinovic, D., & Magliaro, J. (2007). Computer Networks and Globalization. Brock Education Journal, 16(2), Article 2. Https://doi.org/10.26522/brocked.v16i2.84

No existing regulations address these nuanced and especially challenging issues. In addition, regulation currently also fails to ensure compliance where there are gaps in oversight between education and other public sector or commercial fields.

6. The use of AI systems in education is likely to exacerbate the digital divide, both within and between member states.

The digital divide, the gap between those who have access to and can effectively use digital technologies, and those who do not, emerges from differences in socioeconomic status, geographic location, education level, age, and/or physical ability. It can exacerbate existing inequalities by limiting access to information, education, and other essential services. AI-enabled technologies, especially when used in education, are likely to exacerbate this digital divide, both within and between member states.

In particular, AI-enabled technologies are likely to worsen the discriminatory effects of access on economic grounds, for children with disabilities or from communities with minority languages, or those who are otherwise marginalised.[13] This is partly because of both the explicit and hidden costs of the AI-enabled tools, and partly because of their complexity, which although often hidden,[14] impacts greatly on how teachers and learners might critically engage with them.

7. Current policies do not sufficiently protect learners, educators, or education institutions AI-enabled technologies are used in education. Accordingly, a new legal instrument to regulate the use of AI systems in education is needed.

Around the world, wide-scale regulatory provisions that take the protection of democratic participation or the protection of the child and its future self into consideration do not exist. This is especially the case regarding choices made by those in guardian roles in loco parentis (e.g., education institutions) or when assessing the impact of the commercial sector on children's rights. In particular, there is a deficit of regulations specifically focused on promoting and protecting the rights of learners and teachers who engage with AI-enabled educational technologies. We are also lacking regulations designed to protect member states' delivery or oversight of public sector education in a future increasingly impacted by AI.

A recent Council of Europe survey found that few member states have established policies for AI that address education as a special or distinct case.[15] Meanwhile, a UNESCO global survey found that fewer than 10% of schools and universities have developed institutional policies and/or formal guidance concerning the use of AI-enabled technologies in education.[16] The existing regulatory work and evaluation frameworks in some member states should inform the proposed legal instrument, to build on existing expertise, to enable appropriate convergence across all member states, and to ensure that provisions address the protection of democratic participation[17] and the protection of the child and its future self and society. This includes consideration of initiatives underway to integrate responses from the communities that they affect.[18]

---

13. Council of Europe (2019) Two clicks forward and one click back: Report on children with disabilities in the digital environment https://rm.coe.int/two-clicks-forward-and-one-click-back-report-on-children-with-disabili/168098bd0f

14. National Institute of Standards and Technology Interagency or Internal Report 8312 (NIST) Four Principles of Explainable Artificial Intelligence (2021) https://doi.org/10.6028/NIST.IR.8312

15. https://rm.coe.int/ai-coe-survey-report-latest-2753-8190-7209-v-1/1680aec34c

16. UNESCO 2023 survey: Less than 10% of schools and universities have formal guidance on AI https://www.unesco.org/en/articles/unesco-survey-less-10-schools-and-universities-have-formal-guidance-ai

17. Algorithm Watch (2023) ChatGPT and Co: Are AI-driven search engines a threat to democratic elections? https://algorithmwatch.org/en/bing-chat-election-2023/

18. As an example, there is work going on in the city of Barcelona around AI and work on making the edTech approach more democratic and accountable and less BigTech driven. Authorities are working together with the education community and digital rights defenders to be more democratic and offer schools choice and agency. The Catalonia government will join the project in spring 2024. https://ajuntament.barcelona.cat/innovaciodemocratica/en/dd-education-suite-democratic-digitalisation Or a proposed approach in Poland in 2018, for an AI Strategy including education about AI to include skills, competencies and attitudes.

Key issues that need to be addressed in addition to digital competencies, include human rights and child rights, intellectual property, and automation bias, all which must be properly considered when institutions are deciding whether and what AI-enabled tools might be implemented. The result is that teachers, school leadership and policymakers are often unable to assess the emerging risks (for example, to privacy and freedom of thought) when AI-enabled tools are used (e.g., to monitor pupil communications or to automate decision-making). Without clear guidance and governance of tools, it is difficult for key stakeholders to appropriately interpret, compare or assess claims made by the commercial sector. On the other hand, without appropriate information and support, education stakeholders will be unable to benefit from AI-enabled technologies.

Commercial developers also prefer to operate in well-regulated markets, so that they can develop and deploy products that are well-placed to be both unproblematic and profitable. Meanwhile, a key argument often made against regulation is that it will undermine innovation in this fast-moving space. However, there is little evidence for that argument (indeed, the pharmaceutical industry, which is one of the world's most regulated industries, is also one of the most innovative).

Without dedicated protective transparent mechanisms in place, there is a risk that education systems will increasingly depend on AI-enabled technologies that can be proprietary or siloed, often funded by private equity, and over which there is no regulated democratic control. This creates risks to state sovereignty of the content, values, and delivery of core education, and creates new risks to the sustainability of the costs and responsibility of member states' obligations for the provision of education. Without appropriate regulation for their efficacy and safety, AI-enabled technologies could lead to poor, ineffective and unsafe pedagogical practices and disinformation that hinder rather than enhance a learner's intellectual, emotional, and social development.

A clear regulatory eco-system will help member states and companies, all of which would benefit from economies of scale avoiding or minimising duplicate efforts and resources to address similar risks. For example, if an AIED company is awarded some kind of licence to deploy their technology in one member state, they would be able to deploy it in all member states without further financial or opportunity costs. However, currently across Europe, sustainable growth of the AIED and EdTech eco-systems is hindered by fragmented regulatory processes. This has led to almost 65% of European EdTech companies reporting that they have considered leaving Europe to move into other markets (such as North America[19]). Unified regulation across the Council of Europe member states could improve cross-border collaborations, localisation, and AIED's sustainability and reliability.

Some narrow aspects of the emerging EU AI Act (such as the prohibition of emotion detection) will apply to some Council of Europe member states. However, there are many caveats, which means that learners of all ages will remain exposed to harm using technologies that were chosen for them with limited personal choice. In education, the assessment of high-risk AI will have to be more nuanced. Something that is perceived as high risk to one set of learners, may, in fact, be necessary to uphold the right to education of others. For example, AI-enabled eye-tracking technologies may be essential in assisting some disabled learners but also intrusive when creating profiles or enabling social scoring systems.

In addition, while certain products might meet the Act's requirements when they are first deployed in schools, they are likely to be developed over time, thus compromising the Act in ways that go unnoticed. However, there is no duty on schools to monitor the outputs or outcomes of products throughout their lifetime of use in the school from problem formulation

16. European EdTech Alliance, The European EdTech Map: Insight Report 2023
https://static1.squarespace.com/static/5fac2fdb0da84a28cc76b714/t/63bfda44de4b365544ae4b45/1673517650701/EEA+Edtech+Map+Insights+Report+2022.pdf

to decommissioning.[20] Data protection laws are already failing children in particular with regards to automating the human error in algorithms[21] and will be inadequate for the safe use of technologies beginning to be used in educational settings, such as immersive AR (Augmented Reality),VR (Virtual Reality) and haptics.

Meanwhile, self-regulation practices, which are frequently suggested by industry representatives, raise multiple issues, such as the lack of an enforcement process, lack of accountability, and the lack of a clear responsibility to society. The interests of the learner (children and adults) must be a primary consideration in all AI-enabled education activities. However, where incentives are not aligned across the full range of stakeholders (the company, institution, staff, state, and family) the best interests of the learner are compromised: a learner's interests cannot be realised in practice.

Finally, existing regulations also do not offer consistent effective assistance or complaint procedures. Accessible remedies, routes for redress, and effective accountability are all missing for children and young people who experience individual or collective harm to the person or infringements of rights due to the use of AI in educational settings. This is partly due to the power imbalance between children and their families on the one hand, and the commercial sector and the authorities on the other. The rapid deployment of AI-enabled technologies into schools makes it increasingly important to regulate where accountability lies. It must be clear to whom the developers are accountable with transparent methods for addressing and communicating about the embedded technologies and functions.

---

20. ICO (n.d) Monitoring fairness in AI across its lifecycle https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/annex-a-fairness-in-the-ai-lifecycle/

21. Barassi, V. (2020 The Human error in AI and Children's Rights https://childdatacitizen.com/wp-content/uploads/2020/06/The-Human-Error-in-AI-and-Children-Rights_Prof.-Barassi_Response-to-AI-White-Paper-.pdf

3 BACKGROUND

3.1 Artificial Intelligence in Education

> Frequently (e.g. Davies et al. 2020; OECD 2021; Seldon and Abidoye 2018), although rarely with strong evidence (Miao and Holmes 2021a), AI is hailed as a solution to many of education's core problems (e.g. the lack of qualified teachers, student underachievement and the growing achievement gap between rich and poor learners). Nonetheless, this raises the need to consider multiple issues: the aims of using AI in education, where it is used, by whom (by individuals, institutions or industry), how it is operationalised, at what levels (from the single learner to whole classrooms, collaborative networks and national and transnational levels), how it works and so on. (Council of Europe, 2023, p17)

As we have noted, while the use of Artificial Intelligence (AI) in education (AIED) has been a topic of research for over 40 years, there is still a distinct lack of independent evidence at scale for the efficacy and safety of AIED applications.[22] There is not even any agreed way of evaluating or measuring the impact of these AI-enabled tools[23]. Outcomes of AIED might not only shape an individual child's experience of education in the moment, but might also affect their neurological, cognitive, emotional and social development, for life.[24] In fact, AIED tools, the most common being adaptive tutoring systems that claim (without sufficient evidence) to enable new levels of personalisation for learners, are de facto designed to shape the development of human cognition.[25] However, while there are initiatives for regulating AI in general in Europe and elsewhere around the world, there are no frameworks that specifically address key characteristics of education or the human rights, needs and interests of children, young learners and their teachers.

Since AI systems can work unnoticed and at great scale, there is also a real risk of widespread exclusion and discrimination.[26] This is compounded by the fact that children interact with or are impacted by AI systems that are not designed for them, which is something current policies do not adequately address.[27] AIED algorithms are also already being used to automate decisions over access to education and school places for pupils and staff[28], to match teachers to vacant positions[29], and to block students from entering Higher Education based on AI-assigned grades[30]. However, there is a lack of research looking into how much and what types of AI-enabled automation are effective, appropriate, and "good;" or whether these practices are ethical, and whether the human rights of the users are being protected. This involves both those children being targeted and the effects this can have on their future learning experiences.

---

22 Holmes, W. (2023). The Unintended Consequences of Artificial Intelligence and Education. Education International Research. https://www.ei-ie.org/en/item/28115:the-unintended-consequences-of-artificial-intelligence-and-education.

23. Education Services Australia (2023). 'AI in Australian Education Snapshot: Principles, Policy, and Practice': https://www.esa.edu.au/docs/default-source/default-document-library/ai-in-australian-education-snapshot---principles-policy-and-practice_august-2023.pdf. USA Office of Educational Technology (2023). 'Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations': https://www2.ed.gov/documents/ai-report/ai-report.pdf.

24. Council of Europe (2022) AI and Education: A critical view through the lens of human rights, democracy and the rule of law, p37.https://rm.coe.int/artificial-intelligence-and-education-a-critical-view-through-the-lens/1680a886bd, p37.

25. Tuomi, I. The Impact of Artificial Intelligence on Learning, Teaching, and Education. Policies for the future, Eds. Cabrera, M., Vuorikari, R & Punie, Y., EUR 29442 EN, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-97257-7, doi:10.2760/12297, JRC113226

26. UNESCO (2021) AI and education Guidance for policymakers https://unesdoc.unesco.org/ark:/48223/pf0000376709/PDF/376709eng.pdf

27. UNICEF (2021) Policy guidance on AI for children https://www.unicef.org/globalinsight/media/2356/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf

28. Vervloesem, K. Algorithm Watch (2020) In Flanders, an algorithm attempts to make school choice fairer https://automatingsociety.algorithmwatch.org/report2020/belgium/belgium-story/

29. The algorithm that blew up Italy's school system. (2023). AlgorithmWatch. Retrieved 4 August 2023, from https://algorithmwatch.org/en/algorithm-school-system-italy/

30. Fitzgerald, S. (2023). Covid-19 and the International Baccalaureate: A Computer-Assisted Discourse Analysis of #ibscandal. *British Journal of Educational Studies*, 71(2), 129–148. https://doi.org/10.1080/00071005.2022.2056575

3.2 The unique cases of children and education

Children are a special case because of their physical, emotional, and cognitive immaturity. Education also holds a special status because it is essential for the development of informed, responsible citizens who actively participate in democratic processes and address societal challenges. It is also a special case because most often the participation of children is mandatory and decisions about which tools to use and content to access are made by institutions or other bodies and not by the individuals (teachers, students or parents). The right of access to education is a fundamental human right which can be challenged by AIED systems. It becomes questionable, for example, whether AIED tools can be impartial and unbiased, affording all learners access to education. Certain learners and educators could be disadvantaged and their learning and teaching limited or stopped, either intentionally or unintentionally, by biased or even poorly designed systems.

3.2.1 Human rights and child protection

The Universal Declaration of Human Rights, Article 26, states that everyone has a right to education, and that parents have a prior right to choose the kind of education that shall be given to their children (Article 26(3)). In fact, the UNCRC[31,32] goes further. It also states that children are entitled to protection from economic exploitation (Article 32), raises the key issue of privacy (Article 12), and recognises (General Comment no. 25 (IV)(1)) that parents are duty bearers in adjudicating children's enjoyment of their rights to privacy and in the realisation of their rights in the digital environment.

Rights under emerging threat from AIED include the right to freedom of thought. An absolute right enshrined in Article 9 of the Convention, the right to deeply hold any belief (whether religious or not) is absolute and unconditional; the State cannot interfere with it, for instance by dictating what a person believes or taking coercive steps to make him change their beliefs (Ivanova v. Bulgaria, 2007, § 79; Mockutė v. Lithuania, 2018, § 119)) AI is used to nudge opinion, beliefs and behaviour in educational settings. AI is already used in school safeguarding, fraud and plagiarism detection tools and technology used to steer intent from thoughts for 'safety and wellbeing' from written words, even those that are deleted by the student and not shared with any other person. Emerging tools are being marketed at schools to measure brain activity and infer attentiveness. Regulations fail today to address educational technology designed to influence opinions, or that interfere with the right to keep our thoughts private, the right to keep our thoughts free from manipulation, and the right not to be penalised for our thoughts.

In educational settings the role of guardian can fall to educators or even institutions in loco parentis as they navigate the digital environment and make decisions on behalf of children (e.g., which services will be used, and which data will those services collect). Sometimes choices made in the best interests of the institution, educator or even parent may fail to adequately align with the child's views or consider future implications for the child (e.g., the posting images of learners online or in school marketing)[33]. It is not enough to just think about protecting the child in its current situation, especially in a digital environment. Instead, it is important to also protect the child's future self. This means understanding the impact that interactions with AIED can have on a child's future and their experiences of that future. This is not properly addressed in other regulation or data protection law.

---

31. The UN declaration of Human Rights https://www.un.org/en/about-us/universal-declaration-of-human-rights

32. The UN Convention on the Rights of the Child (UNCRC) https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-child

33. Cannataci, J. A. (2021). Artificial intelligence and privacy, and children's privacy Report of the UN Special Rapporteur on the right to privacy. (A/HRC/46/37). https://documents-dds-ny.un.org/doc/UNDOC/GEN/G21/015/65/PDF/G2101565.pdf?OpenElement

One approach, The Montreal Declaration,[34] does seek to address the future, approaching it from three perspectives: the effect on life (including emotional and psychological well-being), the effect on reputation, and the detrimental effects on children's future (including their democratic participation). It is important to note that children's reputations not only influence their interpersonal relationships, but also may have an impact on their ability to access services and employment as they enter adulthood.[35]

Current regulations do not take the protection of the child and its future self adequately into consideration, especially regarding choices made by those in guardian roles (e.g., education institutions). This is also true when education authorities determine how they can monitor and address their obligations centred on the impact of the business sector on children's rights[36] (UN General comment No.16., 2013).[37]

For example, the right to privacy and protection of family life is also a right to the protection of reputation that is inadequately considered in educational settings where inferences from automated decision making may be carried forward throughout a child's school life in ways that were not previously possible. This is especially important when dealing with products that are offered 'for free' but where learners, educators and institutions effectively become the product of their own work. Their data is used, covertly and without the knowledge of the user, to train the AIED system. For example, AIED users are often required to input the product of their own efforts into the system, the provider of which has a clear commercial interest in owning and using this data. A common methodology for technology impact assessment and enforcement is vital in ensuring that the impact of AIED technologies on human rights, cognitive development and pedagogy can be thoroughly assessed before implementation.

3.2.2 Consent

Considering the general recommendations for competence building,[38] regarding what AI is and what it is capable of, it is unrealistic to assume that minors or indeed the majority of students of any age can understand or make fully informed consent decisions regarding the potentially far-reaching consequences of personal data being used to shape their cognitive development, in AIED for training models, facial recognition datasets, and other possibilities.

While not a feature of the Council of Europe Convention 108,[39] for the protection of individuals with regard to the processing of personal data, the General Data Protection Regulation (GDPR[40]) Article 8 introduced the concept of a hard boundary of age (sixteen, but that can be modified to thirteen) regarding the lawful basis of consent for data processing which dictates when parental consent is no longer required for children's data processing.

---

34. The Montreal Declaration for a Responsible Development of Artificial Intelligence was announced on November 3, 2017 at the conclusion of the Forum on the Socially Responsible Development of AI, held at the Palais des congrès de Montréal (EN)https://declarationmontreal-iaresponsable.com/wp-content/uploads/2023/04/UdeM_Decl_IA-Resp_LA-Declaration-ENG_WEB_09-07-19.pdf

35. Carly Nyst, Amaya Gorostiaga, and Patrick Geary. 2018. Industry Toolkit: Children's Online Privacy and Freedom of Expression. UNICEF. Retrieved from https://web.archive.org/web/20210308214931/https://sites.unicef.org/csr/files/UNICEF_Childrens_Online_Privacy_and_Freedom_of_Expression(1).pdf

36. General comment No. 16 (2013) on State obligations regarding the impact of the business sector on children's rights https://www.refworld.org/docid/51ef9cd24.html

37. General comment No. 16 (2013) on State obligations regarding the impact of the business sector on children's rights https://www.refworld.org/legal/general/crc/2013/en/102811

38. Recommendation CM/Rec(2019)10 of the Committee of Ministers to member States on developing and promoting digital citizenship education https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=090000168098de08

39. Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (CETS 108) https://www.coe.int/en/web/data-protection/convention108-and-protocol https://www.coe.int/en/web/data-protection/convention108-and-protocol

40. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) https://eur-lex.europa.eu/eli/reg/2016/679/oj

The notion of age appropriateness recognises that, as children acquire enhanced competencies, there is a diminishing need for protection and a greater capacity for them to take responsibility for decisions affecting their lives.[41] However, a misinterpretation of a broad "digital age of consent," can be used to justify the removal of parental or guardian's involvement in decision-making for older children, contrary to the UNCRC.

Consent only remains valid if informed and freely given, both of which are rarely achievable for minors in the context of educational settings, especially when related to the use of complex technologies. Consent is only valid if it can be withdrawn as easily as it is given. This is impossible once personal data has been incorporated into training datasets for AI systems or used in product development. Mechanisms need to be in place for a child to have their rights managed at various levels in a setting where they are effectively disempowered, and when their agency is compromised by AI-enabled technologies.[42]

Furthermore, examples are emerging of school laws being used to circumvent critical data, consent, and protection issues in order to implement AIED systems. **For example, to implement an AI tool which had received notable criticism, one state in Germany claimed that, so long as the school determines that a teaching and learning tool is mandatory, then no consent of parents or students is necessary**.[43] It becomes important to ensure that there are protections in place for students where, for example, service contracts necessitate AIED use, undermining individual consent.

3.2.3 Procurement and accountability

Procurement, purchasing and implementation practices for education technology tools differ between countries. This becomes increasingly important given the fact that many AIED companies work cross-borders, and the technologies themselves are available online. Currently it remains unclear to whom the companies are accountable when we talk of fairness, safety, and transparency in AIED. Is the responsibility to the school as purchaser, the parent as legal guardian, or to the child as user? It will be important to determine adequate accountability for all stakeholders and, particularly for those whose use of a tool is made mandatory or where consent is diminished.

There is also a rapid development of ecosystems of AIED applications built based on other applications or embedded in existing products through plugins. As a result, a number of different AIED products can co-exist in one product, and each may influence the other. This will make it increasingly difficult but important to determine where accountability for child protection lies and who within this chain of development has responsibility for specific AIED use cases. Furthermore, there are legal uncertainties, for example, whether content created using generative AI tools are the responsibility of the user or the tool's maker or provider, and who should be liable for any infringement.[44] Companies are often based in different locations from their users and the storage location of source data which may bring added complexity to how educators or learners understand copyright.[45]

---

41. Lansdown, G.,(. (2005). The Evolving Capacities of the Child, Innocenti Insights, no. 11. UNICEF. https://www.unicef-irc.org/publications/pdf/evolving-eng.pdf

42. Tuomi, I. (2023). *Measuring the impact of large language models on agency development. In review.* Tuomi, I., Cachia, R., & Villar-Onrubia, D. (2023). *On the futures of technology in education: Emerging trends and policy challenges (JRC Science for Policy Reports). European Commission. In press.*

43. Datenschützer (January 2024) tr. Data protection office: Parents do not have to agree to AI in school | Eltern müssen KI in der Schule nicht zustimmen, Table Media, Bildungstable https://table.media/bildung/news/datenschuetzer-eltern-muessen-ki-in-der-schule-nicht-zustimmen/

44. Walsh, K. (2023). How We Think About Copyright and AI Art. Electronic Frontier Foundation. https://www.eff.org/deeplinks/2023/04/how-we-think-about-copyright-and-ai-art-0

45. Noting that the Digital Markets Act is without prejudice to the Copyright Directive and its transposition in the Member States, with regards to the EU DSM Directive here are considerations for copyright exceptions and using data for machine learning in the EU. The overlap with education in terms of the digital delivery of lectures and teaching and certainly implications for learners' work are less frequently considered in IP discussion than the content of written papers

The integration of AI-enabled technologies and functions into either existing education technology products (EdTech) or standard software such as Microsoft Office and web browsers that are used in educational settings is problematic. Especially when a deactivation of these AI-enabled technologies and functions is not made possible and there is no requirement to transparently inform users of the fact AI is being used within these tools. Such instances of embedded AI systems can provide direct communication channels and data transfer to the AI providers without the end-users (e.g., pupils or students) being properly informed. Service level agreements do not always exist, and contracts often contain small print that limits user rights or contracts are signed at a state or national level without the possibility of individual review.

3.2.4 Commercial influence and stability

The potential influence of commercial entities in AIED tools and their potential for exploiting children's engagement for profit are concerning. Indeed, "AI's true pedagogical use is still questionable, but its political economy is much clearer"[46]. Legal uncertainties also exist when companies use learning analytics data for their product development, particularly regarding due-diligence obligations. AIED tools are granted large amounts of control and influence over the provision, delivery, and access to education for millions of children, but are outside of any democratic control. In other words, decisions are made by individuals or individual corporations for the use by many, and generally without broader consultation. Policymakers must also pay attention to anti-competitive practices, including the effect of proprietary freeware on innovation and SME growth, vertical integration in the market, control of distribution and devices, and global political and policy influence.

Commercial AIED tools, like other EdTech systems, can be affected by regulatory decisions (e.g., when a product is deemed to not be compliant with the GDPR). However, without a centralised manner of addressing this with requirements for grace periods, safe data transfer or assessing budgeting issues for replacement services, this can create instability for educational institutions, educators and learners and can have wide-ranging impact within the ecosystem.

3.2.5 Academic integrity

Questions of authenticity, plagiarism, IP (Intellectual Property) protection and appropriate acknowledgement of original sources or indeed the definition of originality do not currently find satisfactory answers in existing recommendations (e.g., The Council of Europe Recommendation CM/Rec(2022)18 to member States on countering education fraud[47]). The UK Higher Education admissions body (UCAS), for example, has stated that, if their anti-plagiarism software detects elements of a personal statement that may have been generated by AI, it may harm an applicant's prospects for admission.[48] Yet, there are already many pieces of research that question the accuracy of detecting generative AI outputs.[49] In fact, OpenAI

https://mse.dlapiper.com/post/102ivrx/training-ai-models-content-copyright-and-the-eu-and-uk-tdm-exceptions The US Copyright Office position may differ: https://www.copyright.gov/ai/?ref=maginative.com

46. Veale, M. (2023) Schools must resist big EdTech – but it won't be easy https://educationdatafutures.digitalfuturescommission.org.uk/essays/competing-interests-in-education-data/schools-must-resist-big-edtech

47. The Recommendation CM/Rec(2022)18 on countering education fraud addresses the need for a common European approach to ethics, integrity and transparency in education. https://rm.coe.int/ok-prems-137222-gbr-2512-cmrec-2022-18-et-expose-motifs-a5-web-1-/1680a96147

48. UCAS (2023) A guide to using AI and ChatGPT with your personal statement https://www.ucas.com/undergraduate/applying-university/writing-your-personal-statement/guide-using-ai-and-chatgpt-your-personal-statement

49. Dalalah, D., & Dalalah, O. M. A. (2023). The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT. The International Journal of Management Education, 21(2), 100822. https://doi.org/10.1016/j.ijme.2023.100822

clearly states that AI detectors do not work[50] and withdrew its own AI-enabled technology aimed to distinguish between AI-written and human-written text due to its low rate of accuracy.[51] Current education and regulatory frameworks also do not adequately address the challenges for staff, when innocent students are wrongly accused by one computer-driven black box, claiming that the student has cheated using another.

3.2.6 Personalised learning and social scoring

Many claims of benefits to be found in AIED focus on the promise of personalised education for all[52] or the ability to increase student engagement[53]. However, there is only weak evidence to support the claim that AI-enabled technologies genuinely personalise learning or even that personalised learning is beneficial to learners[54] as learning must also be collaborative, emotional, and social. This has been reaffirmed in the most recent PISA (Programme for International Student Assessment) results, with student success being correlated with the connection learners had with their teachers and the support they received from their social environments.[55]

In addition, to enable so-called personalised learning and teaching pathways, many AIED tools must, by default, carry out some degree of social scoring (e.g., comparing behaviours, characteristics and outcomes of learners and educators). This is in direct contrast to both the proposal from the High-Level Expert Working Group on Artificial Intelligence that children should be ensured a free unmonitored space of development[56], and the European Union AI Act (see 3.1.1. below).

3.2.7 Further areas of high risk specific to education

It may not be possible to adequately define what 'high risk' use cases are using the current approach outlined in the EU AI Act where risks are unevenly distributed across a population. This becomes clear in educational contexts. For example, while it is argued that biometric technologies (such as AI-enabled identity verification) may be important to guarantee access to education, if based on facial recognition, keystroke tracking, or motion detection, this may also create a barrier for disabled learners (leaving aside issues of unregulated surveillance). How should teachers therefore assess differentiated risk without discrimination if an AIED tool is to be equitably deployed a whole class of learners? In fact, in the education environment, something that is perceived as high risk to one set of learners may be necessary to uphold the right to access education of others. For example, AI-enabled eye-tracking technologies may be intrusive when creating profiles or enabling social scoring systems but essential in assisting disabled learners. Other areas of contention include e-proctoring systems (AI-enabled technologies designed to maintain academic integrity during online exams), which appear only to meet the wants of institutions, are less ready to address the needs of users, and have

---

50. OpenAI How can educators respond to students presenting AI-generated content as their own? https://help.openai.com/en/articles/8313351-how-can-educators-respond-to-students-presenting-ai-generated-content-as-their-own

51. OpenAI (2023) As of July 20, 2023, the AI classifier is no longer available due to its low rate of accuracy. https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text

52. The Atlantic (2017) Artificial Intelligence Promises a Personalized Education for All https://www.theatlantic.com/sponsored/vmware-2017/personalized-education/1667/

53. Zia, T. (2023) Transforming Education: AI-Powered Personalized Learning Revolution https://www.techopedia.com/transforming-education-ai-powered-personalized-learning-revolution

54. Herold, B. (2017) Education Week The Case(s) Against Personalized Learning https://www.edweek.org/technology/the-cases-against-personalized-learning/2017/11

55. OECD PISA https://www.oecd-ilibrary.org/education/pisa-2022-results-volume-i_53f23881-en

56. EU HLEG-AI (2019) Policy and investment recommendations for trustworthy Artificial Intelligence https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai

exhibited negative consequences for student mental health,[57] automated discrimination, and large data breaches[58].

It is important to establish whether a risk is to be considered higher if the 'harm' affects a large number of students only a small amount, or only a small number of students a large amount. As has been noted, the right to education is indispensable to the furtherance of human rights, however, there is currently no consistent approach to the obligations of educational settings to use AIED systems that uphold the language of teaching or minority languages (important given that many AIED are only in English), or that do not detrimentally affect parents' understanding of the tools used by pupils for homework.

There are many other examples in which a child's human development and future may be affected by AI and algorithms in educational settings, beyond their use in teaching and learning, which data regulations do not properly address. These include:

- tools that claim to monitor children's behaviour, thought, mood and emotions and suggest measurements of engagement and focus;[59]

- school safety technology that uses AI to automate surveillance of children's digital activity and generate risk profiles suggesting self-harm, harms to others or indicators of interests in terrorism and extremism;[60]

- children's image scraping from school websites and their AI-assisted abuse at scale;[61] advertising connected with AI-enabled products;[62] and

- interactions with robots in school premises that may reinforce social norms and influence children's values.[63]

3.3 Regulation

There have been several calls for further regulation of AI in general, and with specific reference to children. For example, one of the key conclusions of the Rome Strategy (2022-2027), as referenced in the Council of Europe Strategy for the Rights of the Child[64], is *"to address the need for legally binding frameworks for AI specifically used by children or for systems that affect children up to the age of 18"*. Meanwhile, the Recommendation on Digital Citizenship Education,[65] adopted by the Council of Europe's Committee of Ministers in 2019, also looked at AI systems being used in education calling for an awareness of the strengths, weaknesses, benefits, and challenges of using AI in learning and schools. The need to protect children's privacy in digital environments is clear. And yet, there are many facets to AIED tools, which are not explicitly defined and where a clear path to regulation does not yet exist. The CAI Framework convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of

57. Retta, M. (2020) Teen Vogue Exam Surveillance Tools Monitor, Record Students During Tests
ttps://www.teenvogue.com/story/exam-surveillance-tools-remote-learning

58. ProctorU confirms data breach after database leaked online (2020)
https://www.bleepingcomputer.com/news/security/proctoru-confirms-data-breach-after-database-leaked-online/

59. AV magazine (2020) viewsonic brings ai-based mood reading to bett
https://www.avinteractive.com/news/products/viewsonic-brings-ai-based-mood-reading-to-bett-28-01-2020/

60. Center for Democracy and Technology. (2021). Student activity monitoring: Polling and research. https://cdt.org/wp-content/uploads/2021/09/Student-Activity-Monitoring-Software-Polling-Research-Slides.pdf

61. Council of Europe. (2001). Explanatory Report for the Convention on Cybercrime. (2001) The Budapest Convention (ETS No. 185). https://rm.coe.int/16800cce5b

62. Council of Europe. (2001). Explanatory Report for the Convention on Cybercrime. (2001) The Budapest Convention (ETS No. 185). https://rm.coe.int/16800cce5b

63. Schiebinger, L., Klinge, I., Sánchez de Madariaga, I., Paik, H. Y., Schraudner, M., and Stefanick, M. (Eds.) (2011-2021). Gendered Innovations in Science, Health & Medicine, Engineering and Environment
https://genderedinnovations.stanford.edu/case-studies/domesticrobots.html#tabs-2

64. Council of Europe Strategy for the Rights of the Child (2022-2027) https://www.coe.int/en/web/children/strategy-for-the-rights-of-the-child

65. Recommendation CM/Rec(2019)10 of the Committee of Ministers to member States on developing and promoting digital citizenship education, https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=090000168098de08.

Law (in progress), for example, does not specifically mention the use of AI in educational settings (the only reference to 'education' is in the context of training AI engineers).

Globally, there are more than 300 policy initiatives in more than 60 countries which aim to govern AI.[66] Education, however, is predominantly only mentioned as part of larger, national AI strategies and there is a paucity of AIED strategies or policies across Europe. In fact, in a recent Council of Europe study of member states, only one in five said that the use of AI in education was regulated while more than half of member states did not respond to the question.[67] Furthermore, while many existing initiatives deal with the application of AI in the context of how it is applied at the implementation stage, they do not address the systemic or structural implications for the education sector, or its political context and capital at local, national, and international levels.[68] In short, addressing how the full range of children's human rights and their future lives may be affected by AIED tools is not yet adequately covered with existing policy. When it comes to children, who have a right to quality education, it is important that an anticipatory rather than a reactive approach underpins risk assessment.

3.3.1 Government regulation

There are various instruments that are being developed nationally and internationally to provide general governance over AI use covering all market sectors. However, this delayed rush for governments and ministries to focus on AI and national issues means that there is less focus on consistent international alignment. Globally, countries are dealing with AI regulation in different ways. In the USA, there has been a focus on a centralised mandate for action (an executive order), which outlines the role that AI will play in important areas of society. This builds on the AI Bill of Rights, and the USA industry agreement[69]. Whilst education is mentioned, there is no detail regarding what this will entail. In the UK, there is no intention to develop new regulation, as the belief is that the impact of AI is covered by existing legislation (which, as we have shown, is often not the case). Other examples of national strategies are from China, India, Germany, and Malta.

Meanwhile, it has been noted that "*if resources related to AI are concentrated in a specific country, we must not have a society where unfair data collection and infringement of sovereignty are performed under that country's dominant position*."[70] In other words, developing cross-border approaches are increasingly important. The AI Safety Summit 2023, held in the UK, brought together international governments and AI companies together with a much smaller number of research experts and civil society to discuss risks and potential mitigation strategies, which might be addressed through international cooperation. The summit resulted in the Bletchley Declaration[71], which outlined a commitment to building shared scientific, evidence-based understanding of impact and to developing risk-based policies in respective national contexts. However, it is important to note that no AI and education experts or NGOs were invited to participate in that summit, which means educational issues are barely mentioned in the declaration.

---

66. OECD repository of over 1000 AI policy initiatives from 69 countries, territories and the EU
https://oecd.ai/en/dashboards/overview

67. https://rm.coe.int/ai-coe-survey-report-latest-2753-8190-7209-v-1/1680aec34c

68. Veale, M. (2023) Schools must resist big EdTech – but it won't be easy
https://educationdatafutures.digitalfuturescommission.org.uk/essays/competing-interests-in-education-data/schools-must-resist-big-edtech

69. Blueprint for an AI Bill of rights https://www.whitehouse.gov/ostp/ai-bill-of-rights/

70. UNICEF Policy Guidance on AI (2021) https://www.unicef.org/globalinsight/media/2356/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf

71. The Bletchley Declaration by Countries Attending the UK AI Safety Summit, 1-2 November 2023
https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023.

European Union AI Act

The EU AI Act, a landmark law for Artificial Intelligence, was finalised by the EU Council and Parliament in early December 2023.[72] This Act outlines responsibilities and ways for AI providers and users to address issues. Its rules vary depending on the potential risk associated with AI usage.

In the field of education, AI raises ethical concerns, especially regarding influencing or manipulating human cognition. As a result, some AI applications used in education are now considered 'high risk'. This classification was further defined following amendments by the European Parliament on June 14, 2023, to include career guidance AI applications, applications used to determine access or materially influence admissions decisions, and applications used for testing appropriate levels of education or monitoring behaviour during testing.

The reason behind this classification is that "AI systems used in education.... should be considered high-risk, since they may determine the educational and professional course of a person's life… When improperly designed and used, such systems may violate the right to education and training as well as the right not to be discriminated against and perpetuate historical patterns of discrimination."[73]

The amended article 4 (214) also explicitly requires member states to promote AI literacy and ensure sufficient knowledge and skills about AI systems and their functions including the different types of products and uses, their risks and benefits. The EU AI Act also prohibits specific AI applications. This includes AI that recognises emotions, real-time biometric systems, systems predicting criminal behaviour, and those that assign social scores to individuals.

However, the majority of educational settings will fall outside the definition of public spaces (9). This means that much of what is restricted in public spaces does not apply and various exemptions within the education specific text negates its purposes. In education specifically, the ban on emotion recognition (Article 5(1)(dc)) has an exception for health and safety reasons (II)(3), which is exactly how much of the technology using AI is marketed to serve learners through mental health support or security purposes. There are serious concerns about the underlying scientific basis for claims made around mood and emotional 'health' detection in on-screen behavioural monitoring of language, or state of mind, or to 'predict' students' interests in terrorism and extremism. This exception in the Act leaves children vulnerable to exposure from some of the most intrusive and high-risk technology to child development, in settings where they cannot refuse to use it. The 'risk-based approach' leaves learners and staff exposed to subjective decision making and inconsistent application of the law. Facial detection and facial recognition may either banned or labelled high risk in some sectors by the Act for law enforcement, but in practice it is widely used across educational settings in the detection of 'fraud' in examinations. While the Act (35) addresses improper use the classification as high risk suggests some acceptable level of risk, it is unclear how educational settings should assess this for compatibility with students' human dignity, or the ECHR Article 14 and a ban on discrimination. The Act points educational settings towards 'how' they can identify and use high risk tools but does not offer educational settings robust rules compatible with the ECHR rights of the child or if' they should use such tools with minors of vulnerabilities and capacities at all.

---

72. EU Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world (2023/2024) https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/

73. Maynard, a. (noting written mid-2023 pre-finalised text) What does EU Artificial Intelligence regulation mean for ai in education? https://futureofbeinghuman.com/p/eu-ai-act-and-education

GDPR / Data Protection and AIED

The European Data Protection Board has formed a task force to address questions of lawfulness and cross-border cooperation on generative AI such as ChatGPT and decisions are expected to follow.[74] Data protection may enable the use of data in many cases within the law, but in ways that can fail to uphold fundamental human rights to privacy, or that are detrimental to human dignity and the promotion of young people's flourishing. For example, in the use of biometric technology in schools, there is inconsistent practice reliant on current EU data protection law. Whilst some countries have made court or regulatory decisions governing this[75.], there is fundamental criticism, for example, that "the law does not provide the legal guardrails necessary to ensure that Facial Recognition Technology (FRT) is developed and deployed in ways that uphold basic human rights."[76]

3.3.2 Principles

Hundreds of voluntary or non-binding principles documents have been created aiming to guide AI strategy and policy. A study on rights-based approaches to AI, however, found that, whilst 64% of all the gathered principle frameworks made a reference to specific human-rights related documents, only five documents actually employed a human rights framework.[77] Both public concern and a lack of research or experience regarding the societal effects of AI have led to a strong focus on ethical AI principles, often at the sake of all other potentially important topics.[78]

There is also a distinct lack of frameworks specifically addressing education. The ones that do, largely focus on developing competences for AI in teaching and learning[79], AI literacy including academic integrity[80], and socio-developmental ethics[81]. Importantly, a review of national AI policy strategies related to the role of education and ethics indicates that "these documents tend to focus on Education for AI (i.e., training) rather than on AI for Education (AIED)".[82] Overall, there is a distinct lack of resources with a focus on education-specific, rights-based challenges such as child's rights, child and youth development, and societal aspects. The fact that students travel across geographical boundaries to study and work may cause confusion in practice where national acceptable and lawful practices diverge.

3.3.3 Self-regulation

Self-regulation and self-imposed guidelines or frameworks can be used to convince legislators that stakeholders can self-govern and that specific legal instruments are not necessary.[83] Following this strategy, AI providers as well as their customers and users of AI technologies

74. Sterling, T. (2023). European privacy watchdog creates ChatGPT task force. *Reuters.* https://www.reuters.com/technology/european-data-protection-board-discussing-ai-policy-thursday-meeting-2023-04-13/

75. Swedish Data Protection Authority case against Skellefteå Municipality, Secondary Education Board DI-2019-2221 https://gdprhub.eu/index.php?title=KamR_Stockholm_-_Case_No._5888-20

76. Davis, P. N., Perry, L., & Santow, P. E. (2022). Facial recognition technology: Towards a model law. Human Technology Institute, The University of Technology Sydney. https://www.uts.edu.au/human-technology-institute/projects/facial-recognition-technology-towards-model-law

77. Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., and Srikumar, M. (2020). Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI. Berkman Klein Center Research Publication No. 2020-1, Available at SSRN: https://ssrn.com/abstract=3518482 or http://dx.doi.org/10.2139/ssrn.3518482

78. Holmes, W., & Porayska-Pomsta, K. (Eds.). (2023). *The Ethics of AI in Education. Practices, Challenges, and Debates.* Routledge. https://www.routledge.com/The-Ethics-of-Artificial-Intelligence-in-Education-Practices-Challenges/Holmes-Porayska-Pomsta/p/book/9780367349721

79. European Union (2022) Final report of the Commission expert group on Artificial Intelligence and data in education and training https://op.europa.eu/en/publication-detail/-/publication/7f64223f-540d-11ed-92ed-01aa75ed71a1/language-en/format-PDF/source-search

80. Russell Group new principles on use of AI in education (2023) https://russellgroup.ac.uk/news/new-principles-on-use-of-ai-in-education/

81. e.g. Tuomi, I. (2023). A framework for socio-developmental ethics in educational AI. *Proceedings of the 56th Hawaii International Conference on System Sciences, 6208–6217. https://hdl.handle.net/10125/103386*

82. Schiff, D. Education for AI, not AI for Education: The Role of Education and Ethics in National AI Policy Strategies. Int J Artif Intell Educ 32, 527–563 (2022). https://doi.org/10.1007/s40593-021-00270-2

83. Calo, R. (2017). Artificial intelligence policy: a primer and roadmap. SSRN Journal, 1–28.

can claim to address and be aware of ethical questions and concerns but without any implications and effects on the AI applications and practices. This raises multiple issues, such as the lack of an enforcement process, lack of accountability, and the lack of a clear responsibility to society.

UNESCO calls for Governments and ministries of education to develop their own capacities for review and validation of AI to reduce any reliance on industry self-regulation[84]. An analysis of 22 major ethics guidelines further highlights a common issue in that "AI ethics—or ethics in general—lacks mechanisms to reinforce its own normative claims" and that principle frameworks "are rather weak and pose no eminent threat"[85] to any of the stakeholders.

With the precedent being set this year by the ability of OpenAI's leadership to overthrow the safeguards of their advisory board[86], it is essential to ensure safeguard structures exist, which can go beyond the benefits of shareholders or individual companies and look to the requirements of learners, educators, and the societies within which they find themselves.

3.3.4 Standards

Standards will play a vital role in supporting compliance to any regulations by defining concrete technical requirements to adhere to. However, even with standards in place, mechanisms must still be developed to help all stakeholders ensure they are compliant, and the participation must be ensured by those most interested in protecting fundamental human rights and the public interest[87].

General AI standards rarely address education-specific issues which can increase issues of access and inclusion in educational settings and the devaluation of regional or minority languages and their protection and promotion contributing to the building of a Europe based on democracy and cultural diversity.[88] Design and technology standards are also not obliged to uphold SDG (Sustainable Development Goals) 4 or to ensure that EdTech is inclusive, equitable and open to all, which could unintentionally exacerbate issues of unfairness, exclusion, and poor pedagogical practice). Whilst there are some AI standards under development[89] it is important to note that standards like these are voluntary, non-binding instruments and often exist behind a paywall.

3.3.5 Wide-ranging ecosystem support

In the public sector, numerous papers and frameworks from UNESCO, UNICEF, OECD, and the European Commission explore different aspects of AIED implementation and policy. These range from global analyses of current policy and AI curricula, high-level policy recommendations, technical recommendations for AI implementation, and teacher training playbooks.

Industry representatives from Google, Microsoft and OpenAI (among others) have called for regulation beyond self-regulation practices. In April 2023, hundreds of Industry initiatives, and researchers, including many affiliated with the Future of Life Institute, called for a pause in the

---

84. UNESCO (2021) AI and education: guidance for policy-makers https://unesdoc.unesco.org/ark:/48223/pf0000376709

85. Hagendorff, T. The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds & Machines* 30, 99–120 (2020). https://doi.org/10.1007/s11023-020-09517-8

86. Dave, P. (2023). How OpenAI's Bizarre Structure Gave 4 People the Power to Fire Sam Altman. Wired. https://www.wired.com/story/openai-bizarre-structure-4-people-the-power-to-fire-sam-altman/

87. McFadden et al.

88. The European Charter for Regional or Minority Languages https://www.coe.int/en/web/european-charter-regional-or-minority-languages/about-the-charter Together with the Framework Convention for the Protection of National Minorities it constitutes the Council of Europe's commitment to the protection of national minorities.

89. For example, ISO (the International Standardization Organization) in sub-committee JTC1/SC42 https://www.iso.org/committee/6794475.html

development of AI technologies as *"Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable"[90]*.

A need for better understanding and greater literacy around AI, focusing not only on technology but also on its funders and their values, intended and unintended human and societal impacts has driven several third-sector or civil society efforts[91], which focus on supporting educators and learners in using AI for teaching through guidelines, curricula, training programmes and toolkits[92]. It is important to recognise any industry dependencies, including financial and lobbying activities that may influence these initiatives, and that there may only be an indirect interest in human rights, democracy, social justice, or the expression and development of human agency.

3.3.6 Council of Europe's unique role

Unlike AI ethics frameworks, human rights are enforceable in law, and, therefore, more fitting to govern AI throughout its life cycle.[93] The central values behind the Human Rights based approach is reflected in the draft[94] of the Convention on Artificial Intelligence, Human Rights, Democracy and the Law, which provides a high-level coverage of all key issues related to AI governance.

The Council of Europe's Committee on Artificial Intelligence (CAI)'s Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law (in progress) also takes a risk-based approach to AI.[95]

3.4 Why a legal instrument to regulate the use of AI in education

The Council of Europe international conventions offer a range of significant benefits that contribute to the respect, fulfilment, and promotion and protection of human rights, democratic values, and the rule of law. These benefits are particularly evident in three key areas

1. Harmonisation of approach within a human rights-based framework

2. Facilitation of cooperation across member states and observer states

3. Enhanced legitimacy and credibility through enforcement frameworks

The global landscape of AI regulatory practices governing use is multifaceted. There are hundreds of documents outlining proposed ethical principles for use[96] and several recommendations exploring good regulatory practice in general, but not making specific

---

90. Reuters (2023) Elon Musk and others urge AI pause, citing 'risks to society' https://www.reuters.com/technology/musk-experts-urge-pause-training-ai-systems-that-can-outperform-gpt-4-2023-03-29/ It is to note that authors of one of the papers cited in the letter, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" including Emily M. Bender, Timnit Gebru, and Margaret Mitchell, criticised the letter. Mitchell claimed that "by treating a lot of questionable ideas as a given, the letter asserts a set of priorities and a narrative on AI that benefits the supporters of FLI. Ignoring active harms right now is a privilege that some of us don't have."

91. Access Now (2023) What's not in the EU AI Act https://www.accessnow.org/whats-not-in-the-eu-ai-act-deal/

92. The International Society for Technology in Education (ISTE) provides resources for understanding and implementing AI in schools. The Consortium for School Networking (COSN) emphasises leadership guidelines, training, policy development, and integrating privacy and security measures. Multistakeholder initiatives like the World Economic Forum's AI Governance Alliance and the Association for the Advancement of Artificial Intelligence, along with newer organisations like AIEdu and Code.org's TeachAI, are developing safe AI designs, policies, and curricula.

93. Internet Policy review (Vol.12 Issue 1) referencing Donahoe & Metzker, 2019; McGregor et al., 2019; Yeung et al., 2020; Smuha, 2020; Cobbe et al., 2020 https://policyreview.info/articles/analysis/future-proofing-the-city

94. Council of Europe. (2023). Committee on Artificial Intelligence: Revised zero draft [framework] convention on Artificial Intelligence, human rights, democracy and the rule of law (CAI(2023)01; pp. 1–13) https://rm.coe.int/cai-2023-01-revised-zero-draft-framework-convention-public/1680aa193f.

95. Council of Europe CAI (July 2023) Consolidated working draft of the Framework Convention on Artificial Intelligence, Human Rights, Democracy and The Rule of Law https://rm.coe.int/cai-2023-18-consolidated-working-draft-framework-convention/1680abde66.

96. OECD policy repository https://oecd.ai/en/dashboards/overview/policy

recommendations due to the unique nature of education environments and structures. Alone within European Commission Eurydice Network [97].

3.5 Example possible areas of AI regulation specific to education

3.5.1 Protecting the human-rights of learners before AI-enabled technologies are introduced into classrooms

Existing AI regulatory practices seek to address recourse if harm has been experienced and to ban large issues spanning multiple sectors. But they do not look at stopping human rights violations of learners before AI-enabled technologies enter learning environments. A key human right for children is to a quality education, whereas many AI-enabled technologies designed for use in education have limited evidence for their efficacy or safety.

Possible regulatory needs:

- Including human rights aspects in principles and enforcement mechanisms.

- Focussing on the human rights of minors in the use of AI-enabled technologies.

- Requiring robust independent evidence at scale for the efficacy and safety of any AI-enabled technology before it can be used in an educational setting.

3.5.2 Mitigating harm from inappropriate materials in AI-enabled technologies

Currently, when safety issues are identified within an AI-system (e.g., when they are prompted to provide instructions on self-harm, harming others, or committing crimes) the connection made within the system may be throttled.[98] A company 's approach in mitigating harms[99] can also have unintended consequences as Google found in the 2024 launch of its image generator Gemini.[100]

Possible regulatory needs:

- The protection of developmentally appropriate environments aligned with the evolving capacities of the child, and the rights of the parent in the exercise by the child of the rights recognized in the Convention.[101]

- The protection of freedom of expression and content accuracy in creative tools

- The successful elimination of illegal materials and consensus of approach in educational tools on inappropriate outputs and influences perpetuating stereotypes in gender, ethnicity and race and biased content creation.

3.5.3 The complex layers of AI-enabled technologies and tools built upon them

Needing a niche product, a teacher might introduce an AI-enabled app to their classroom. However, that app might show unintended bias towards certain learners. When trying to seek recourse, how should schools or learners understand where a problem is if the app is built on

---

97. European Commission Eurydice network https://eurydice.eacea.ec.europa.eu/national-education-systems

98. https://www.lesswrong.com/posts/7fYxxtZqjuYXhBA2D/testing-ways-to-bypass-chatgpt-s-safety-features

99. Commercial case study: This company offers training on AI bias mitigation (Holistic AI) https://www.holisticai.com/blog/technical-resources-bias-mitigation.

100. Time magazine (2024) "Google Left in 'Terrible Bind' by Pulling AI Feature After Right-Wing Backlash" https://time.com/6835975/google-gemini-backlash-bias/

101. CRIN (n.d.) Parental guidance and the evolving capacity of the child (Article 5 of the UNCRC) https://archive.crin.org/en/home/rights/convention/articles/article-5-parental-guidance-and-childs-evolving-capacities.html and UN OHCHR statement on Article 5 (October 2023) https://www.ohchr.org/sites/default/files/documents/hrbodies/crc/statements/CRC-Article-5-statement.pdf

layers of other applications before even getting to the core AI model behind it, and trained on multiple datasets[102], which makes it easy for responsibility to fall between parties[103].

Possible regulatory needs:

- Frameworks defining chains of development and the regulatory responsibility each layer holds or can be held to.
- Regulatory measures for each layer of development.

3.5.4 Mass scale manipulation of learners or teaching practices

Generative AI (such as ChatGPT) can produce manipulative media (text, images, voices etc.) on a large scale. This process can be amplified by tailoring the content to individuals using personal data. Algorithmic manipulation could be significantly amplified by specifically designed models influencing teaching and learning practices.

Possible regulatory needs:

- Provisions to not exploit, interfere with, or manipulate learners or teachers.
- Provisions to ensure alignment with and checks of pedagogically sound practice.

3.5.5 Clarity about how AI-enabled technologies are trained and their intent

Due to the 'black box' nature of much AI and opacity of automated decision making, it is difficult to trace why a particular recommendation was made or why a particular output was generated. Accordingly, information is necessary about the resources and systems with which the AI-enabled technology was trained, in order to empower sound decision-making about trustworthiness and applicability for use. The educational purpose and intent of the technology should be stated, transparent, verifiable, and checked.

Possible regulatory needs:

- Measures outlining transparency requirements for training and development of AI-enabled technologies and their data outputs for use in education.
- Provisions for transparency of purpose, intent and, where applicable, examples of unintended consequences.

---

102. Thompson, A. D. (2022). What's in my AI? A Comprehensive Analysis of Datasets used to Train GPT-1, GPT-2, GPT-3, GPT-NeoX-20B, Megatron-11B, MT-NLG, and Gopher. https://LifeArchitect.ai/whats-in-my-ai

103. Thomson, A. (2023) https://s10251.pcdn.co/wp-content/uploads/2023/03/2023-Alan-D-Thompson-GPT3-Family-Rev-1.png

4 QUESTIONS FOR DISCUSSION

The following questions may be used as prompts – for readers to consider their own views raised by the paper and to inform discussion.

AI systems used in education

– What mechanisms or infrastructure should be put in place to ensure the **efficacy, safety, and positive pedagogical impact on the eco-system** of the classroom of AI-enabled technologies?

– Whose **relationships,** agency or authority are changed because of the use of AIED, between the **institution, parents, the individual teacher or learner** and introduction of companies into a child's education? Are these changes desirable or do they need mitigation?

– What **level of harms, biases, opacity in automated decisions, and poor educational experiences** is acceptable? What mechanisms and routes for remedy and redress would be appropriate for learners and educators who experience them?

– What if any special use cases should be considered for communities with shared characteristics or rights such as **children with disabilities, minority languages, or the right of parents** to have their children educated in conformity with their religious and philosophical convictions?

– Do the climate and resource implications[104] from the use AI merit special attention in the education sector and educators' responsibility for their own **influence on children's future environment?**

Human rights, democracy and the rule of law

– Which **focus areas** of human rights, democracy and the rule of law should be prioritised in education environments and must be addressed by the proposed legal instrument?

– Is there agreement on the **imperative to protect the human rights of children** in educational settings?

  - Assuming the incentives for business do not always prioritise children's rights, what would be a suitable measure to evaluate an appropriate "balance" between economic drivers for AI with any duties towards children?

  - A risk-based approach might assume some level of discrimination as a result of the application of AIED is acceptable. Is this in line with Article 14 and Article 1 of Protocol No. 12 of the Convention that enshrines the right not to be discriminated against?

  - Should needs and risk be assessed for each individual in a single classroom, year group or school, or should educators treat children as a homogenous group?

– How can the Council ensure that the proposed legal instrument does not **exacerbate the digital divide or undermine access to equitable and quality education**?

– Is there **consensus on what ethical norms** should inform the proposed legal instrument to regulate the use of Artificial Intelligence systems in education?

Regulation

– Given the diversity of educational systems across Council of Europe member states, what **opportunities and challenges will there be when harmonizing the approach**

---

104. Hao, K. (2024) The Atlantic: New data centers are springing up every week. Can the Earth sustain them? https://www.theatlantic.com/technology/archive/2024/03/ai-water-climate-microsoft/677602/

to regulate the use of AI-enabled technologies in education? How soon would member states commit to addressing these challenges and how?

– What will be **the role of all stakeholders** (learners, parents, educators, school leadership, and industry) in ensuring the effective implementation of the proposed legal instrument and how should this be operationalised over what time period?

– How do member states want the Council to **support them to effectively implement the legal instrument, foster collaboration and harmonisation**, and monitor its impact on learners, educators, and learning environments?

– Would member states want to **participate in monitoring of the instrument to share positive and negative experiences of the emerging outcomes on learners** and teaching community to communicate collective learnings from practice?