



OPEN ACCESS

EDITED BY

Athanasios Drigas,
National Centre of Scientific Research
Demokritos, Greece

REVIEWED BY

Victoria Bamicha,
National Centre of Scientific Research
Demokritos, Greece
Cristina-Georgiana Voicu,
Alexandru Ioan Cuza University of Iasi,
Romania

*CORRESPONDENCE

Kostas Karpouzis
✉ karpou@panteion.gr

RECEIVED 25 August 2025

ACCEPTED 22 September 2025

PUBLISHED 10 October 2025

CITATION

Panagopoulou F, Parpoula C and Karpouzis K
(2025) Legal perspectives on AI and the right
to digital literacy in education.
Front. Comput. Sci. 7:1692268.
doi: 10.3389/fcomp.2025.1692268

COPYRIGHT

© 2025 Panagopoulou, Parpoula and
Karpouzis. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Legal perspectives on AI and the right to digital literacy in education

Fereniki Panagopoulou¹, Christina Parpoula² and
Kostas Karpouzis^{3*}

¹Department of Public Administration, Panteion University of Social and Political Sciences, Athens, Greece, ²Department of Psychology, Panteion University of Social and Political Sciences, Athens, Greece, ³Department of Communication, Media and Culture, Panteion University of Social and Political Sciences, Athens, Greece

Introduction: Artificial intelligence (AI) is increasingly becoming a part of educational practice, providing opportunities for personalization and access but also introducing risks to equity, learner autonomy, privacy, and accountability. Focusing on European and Greek contexts, we examine whether a right to digital literacy can be grounded in existing law and how the EU AI Act reshapes duties for educational actors.

Methods: We conduct a legal-doctrinal and policy-analytic study of EU primary/secondary law, the Greek Constitution, and EU Regulations (AI Act), read alongside institutional 'grey' literature (e.g., educator toolkits, national bioethics opinions). We map AI Act recitals and Annex III to concrete governance obligations (fundamental rights impact assessment, transparency, human oversight) and test implications through targeted case vignettes (examinations, admissions, LMS/explainability). Scope is limited to EU/Greece; comparative case law is used selectively to illuminate the normative claims.

Results: First, a defensible right to digital literacy emerges from EU instruments and Greek constitutional provisions on participation in the information society and the mission of education. Second, many AI uses in education (e.g., admissions, outcome evaluation, proctoring) qualify as high-risk under the AI Act, triggering ex-ante and ongoing duties, while emotion recognition in education (absent medical/safety grounds) is effectively off-limits. Third, the vignette analysis shows recurring pressure points, such as bias and disparate impact, opacity in automated decisions, and excessive surveillance, where explainability and meaningful human oversight are necessary to preserve equity, autonomy, and educational quality.

Discussion: We propose an actionable governance agenda for schools and universities: mandatory fundamental-rights impact assessments adapted to educational contexts; explainability criteria as admissibility and accountability thresholds for deployed systems; clear escalation paths and complaint mechanisms; inclusive access measures to narrow the digital divide; and multi-stakeholder oversight that keeps educators central rather than substitutable. Taken together, these measures shift debate from abstract ethics to enforceable legal duties, aligning AI adoption with human-rights values and the educational mission to cultivate critical, responsible citizens.

KEYWORDS

artificial intelligence, digital literacy, equity, human rights, explainable AI, AI ethics, education policy

1 Introduction

Artificial intelligence (AI) is already changing the way in which we learn, work and live, and education is affected by this development (European Commission and Culture, 2022). The introduction of AI applications in education provokes both optimism for innovation and concern about risks; concerns are particularly pronounced, given that minors, meaning individuals whose personalities are shaped through the educational process, are also involved. Therefore, any experimentation should be approached with restraint () and great caution, as potential harm may prove irreversible. This does not mean that the introduction of AI tools in the field of education should be prohibited, but rather that their application ought to be clearly delineated. This must be pursued with a primary focus on enhancing the educational process, without compromising its fundamental objective: the formation of responsible and conscientious citizens and the cultivation of ethical reasoning.

The paper first examines education as a human right, then outlines the relevant legal frameworks, before turning to case studies and policy recommendations. We follow a legal-doctrinal and policy-analytic method of EU and Greek primary law, read alongside institutional gray literature (Benzies et al., 2006), i.e., official guidance, committee opinions, and implementation toolkits, used as interpretive evidence rather than empirical data. In this context, we closely map the relevant provisions in the AI Act (cf. recitals, Annex III) to concrete governance duties (fundamental rights impact assessment, transparency, human oversight) and test their implications through targeted case vignettes (e.g., examinations, admissions, LMS/XAI). To manage bias and recency in gray sources, we take into account provenance (EU/Greek public bodies), issuance purpose (guidance vs. advocacy), and timeliness, drawing on educator guidelines, national bioethics opinions, and state toolkits such as SELFIE. Finally, we delimit scope to EU/Greece and offer no claims about tool efficacy; case law and comparative jurisdictions are treated selectively where they illuminate the normative argument.

In brief, our analysis shows that a defensible right to digital literacy emerges within European and Greek law, and that many educational AI uses fall under the high-risk category of the EU AI Act, triggering obligations such as fundamental-rights impact assessment, transparency, and human oversight; certain uses (e.g., emotion recognition in education, absent medical/safety grounds) are effectively off-limits. In this framework, we contribute a legal-normative framework that connects human-rights values (equity, autonomy, quality, social responsibility) to concrete governance requirements in schools and universities, integrate XAI as a criterion for admissibility and accountability, and distill policy recommendations (AI literacy, explainability scoring, accountability guidance, inclusive access, multi-stakeholder oversight) that are immediately actionable by education authorities. The contribution of this paper is to move debate beyond generic ethics pledges toward enforceable duties grounded in constitutional and EU law.

2 Education as a human right

Education is widely recognized as a fundamental human right, affirmed in international instruments, such as the Universal Declaration of Human Rights (UDHR) and the International Covenant on Economic, Social, and Cultural Rights (ICESCR). Article 26 of the UDHR states that “everyone has the right to education” (UN, 1948), highlighting that education is not a privilege but an essential right for all human beings. This right is further underscored by the ICESCR, which mandates that States “recognize the right of everyone to education” (Vierdag, 1978).

Beyond being a right on its own, education is also viewed as a tool for the realization of other human rights, such as freedom, equality, and dignity. It is often described as “a gateway to the full realization of a wide range of human rights.¹” Education empowers individuals by providing them with the knowledge and skills necessary to participate fully in society, thereby promoting social inclusion, civic participation, and economic opportunities; it also serves to uphold the values of democracy and pluralism, supporting the development of informed and responsible citizens.

2.1 Core values of education

Several core values should be respected within the context of education, which can be understood as both intrinsic and instrumental principles. These values not only define the essence of education but also guide the implementation of educational policies and practices. The following are key values:

2.1.1 Equity and equality

Education must be inclusive and accessible to all, irrespective of an individual's background, gender, ethnicity, or socioeconomic status. The concept of educational equity goes beyond mere access to schooling and addresses the disparities that exist in educational outcomes, aiming to provide every learner with the opportunity to succeed. Equity in education emphasizes tailored support and resources to ensure that marginalized and vulnerable groups are not excluded. This commitment is embodied in the 2030 Agenda for Sustainable Development, which includes a specific target to “ensure equal access for all women and men to affordable and quality technical, vocational, and tertiary education, including university.²” This pursuit of equality is not simply about achieving uniformity in educational provision, but also about considering the different needs of learners to ensure that everyone can reach their full potential.

1 UNESCO, Education for all 2000-2015: achievements and challenges, <https://www.unesco.org/gem-report/en/efa-achievements-challenges>, last accessed: June 11, 2025.

2 Council of Europe, Equal access to technical/vocational and higher education, <https://www.coe.int/en/web/education/4.3-equal-access-to-technical/vocational-and-higher-education>, last accessed: June 11, 2025.

2.1.2 Freedom and autonomy

A core value of education is the promotion of autonomy, freedom of thought, and self-expression: education should encourage individuals to think critically, form their own beliefs, and contribute to public discourse. This is reflected in the European Convention on Human Rights, which upholds the right to education, alongside the right to free expression and intellectual autonomy (Harris et al., 2023). In educational contexts, freedom also means that individuals should be free from coercion or oppression, allowing them to pursue knowledge in an open, diverse, and welcoming environment. The principle of academic freedom, integral to higher education, allows educators to teach, learn, and discuss ideas freely without fear of censorship.

2.1.3 Quality and excellence

While access to education is important to begin with, the right to education should also include the right to quality, relevant, and effective learning experiences. This is of utmost importance, since quality education enables students to develop critical thinking skills, creativity, and the capacity to engage with complex social, economic, and environmental challenges. The UNESCO Framework for Action on Education for Sustainable Development (Unesco, 2015) emphasizes that education should enable transformation, enabling students to become active participants in shaping their futures. In this context, ensuring a high standard of education contributes not only to individual development but also to societal progress.

2.1.4 Solidarity and social responsibility

Education is, in most cases, a communal activity, which cultivates a sense of shared responsibility toward fostering a just, sustainable, and peaceful world. As such, education should instill a sense of solidarity and promote social cohesion, tolerance, and respect for human rights. According to the International Commission on Education for the Twenty-First Century (Delors, 1996), education should not only focus on individual achievement but also on collective responsibility.

2.2 The impact of AI on education

AI has shown the potential to enhance the educational experience (George and Wooden, 2023), but it also raises critical questions regarding the preservation of these core values. AI's transformative capabilities could disrupt traditional educational practices, leading to both positive and negative consequences for equity, autonomy, and the overall quality of education.

2.2.1 AI and equity

AI technologies have the potential to significantly improve access to education. For example, AI-driven platforms can provide personalized learning experiences that adapt to the needs of each student, helping to close achievement gaps by providing additional support where needed (Holstein and Doroudi, 2021). In regions where there is a shortage of qualified educators, AI

can provide virtual tutors or teaching assistants, making education more accessible to underserved populations.

However, AI could also exacerbate existing inequalities if not implemented thoughtfully. One of the central concerns regarding AI in education is the potential for bias in algorithms. AI systems are trained on historical data, and if that data reflects existing biases—such as racial, gender, or socioeconomic biases—the resulting AI tools can perpetuate these biases, disadvantaging certain student groups. To ensure that AI serves the goal of equity, it is vital that AI systems are developed and implemented with attention to fairness and inclusivity, using diverse datasets and transparent methodologies (O'Neil, 2017).

2.2.2 AI and freedom

AI in education could either support or undermine the value of autonomy. On one hand, AI can enhance individual freedom by providing students with more agency in their learning paths. Adaptive learning technologies, for example, allow students to progress at their own pace, enabling a more personalized educational journey. Students can also access a vast range of resources and materials online, giving them greater freedom to choose how and when they engage with learning content.

On the other hand, there are concerns that AI could limit freedom if it is used to excessively monitor, control, or predict students' behaviors. For example, AI systems that track students' activities and performance may lead to an environment where students feel pressured to conform to predefined standards of success. This could undermine the development of independent thought and creativity, as students may prioritize meeting algorithmically defined goals over exploring their own intellectual curiosities. To preserve freedom, it is essential that AI systems in education respect students' rights to privacy and autonomy, ensuring that their use is aligned with educational goals rather than merely optimizing performance (Zuboff, 2023).

2.2.3 AI and quality of education

AI has the potential to enhance the quality of education by offering personalized learning experiences, automating administrative tasks, and providing real-time feedback to students and teachers. AI can help educators identify learning gaps more effectively, enabling timely interventions (e.g., Karpouzis et al., 2024). Furthermore, AI-powered tools such as chatbots or virtual assistants can offer 24/7 support to students, providing guidance and information outside of class hours.

Nevertheless, the use of AI in education also raises concerns about the standardization of education. AI systems could inadvertently narrow the scope of education by prioritizing data-driven, measurable outcomes at the expense of more qualitative, holistic educational goals. For example, an over-reliance on AI-driven assessments may reduce opportunities for students to engage in creative, critical thinking, which are key components of a well-rounded education (Selwyn, 2019). As such, the introduction of AI should be carefully balanced with efforts to maintain diverse educational experiences that nurture various aspects of student development.

2.2.4 AI and solidarity

AI systems in education can foster global solidarity by providing equitable access to high-quality learning materials, especially for students in remote or underserved regions. For example, AI can facilitate language translation services, making educational resources accessible to a wider audience. Additionally, AI can help create networks of learning that bridge geographical and cultural divides, promoting cross-cultural understanding.

However, widespread use of AI could also exacerbate the digital divide if access to technology is not equitably distributed. Students without access to reliable internet connections or the necessary devices could be excluded from AI-enhanced educational opportunities, reinforcing social inequalities. This calls for policies that ensure equal access to AI technologies and the necessary infrastructure, so that all students can benefit from the advancements AI offers.

3 Legislative and advisory framework

There is no specific legal regime governing the use of AI in education. Nevertheless, certain key principles may be inferred from the European legal framework itself, the provisions of the Greek Constitution, Regulation 1689/2024 on AI, and Law No. 4961/2022 on emerging information and communication technologies, strengthening digital governance and other provisions. Guidance is also provided by the opinion of the High-Level Expert Group on Artificial Intelligence (AI HLEG).

More recently, intergovernmental baselines set clearer expectations for AI in schooling. UNESCO's global guidance and competency frameworks (Cukurova et al., 2024) articulate human-centered principles and measurable teacher/student competences, while OECD analyses document how systems are presently governing generative AI, formalizing teacher digital competences, and addressing equity risks (Varsik and Vosberg, 2024); we treat these as policy priors against which our legal-doctrinal claims are assessed.

3.1 The European legal framework

3.1.1 Article 2 of the first additional protocol to the ECHR

In accordance with this article “[n]o person shall be denied the right to education. In the exercise of any functions which it assumes in relation to education and to teaching, the State shall respect the right of parents to ensure such education and teaching in conformity with their own religious and philosophical convictions.” The negative formulation (*no person shall be denied*), as opposed to the initially proposed affirmative one (*every person has the right to education*) gives rise to a dual interpretation. First and foremost, the emphasis lies in ensuring effective access to the educational system provided by the state.³ Secondly, the state is not under a positive obligation to take active measures to guarantee access to education of one's choice, to allow individuals to design

their own educational systems, or to subsidize private education (Tamamidis, 2021). The provision does not impose a duty on the state (Margaritis, 2018) to provide selective education; such matters fall with the state's discretion (Harris et al., 2023). The case law of the European Court of Human Rights (ECtHR) highlights the need for pluralism⁴ in education. This implies that AI systems deployed in the educational context must also foster pluralism—both in terms of the means of instruction (e.g., AI tools) and the content taught (e.g., multicultural perspectives).

3.1.2 Article 14 of the EU charter of fundamental rights

Under Article 14 of the Charter, everyone has the right to education and access to vocational and continuing training. This right includes the opportunity to attend compulsory education free of charge. The wording of Article 14 broadens the scope of the right to education in comparison to the corresponding provision of the ECHR, enshrining—among other things—access to vocational and continuing training, as well as the right to attend compulsory education free of charge. Education is understood as “the acts or processes by which, inter alia, information, knowledge, perceptions, attitudes, values, skills, abilities, or behaviors are transmitted or acquired.”⁵ In a modern digital state, continuing training necessarily includes digital education, as the very standards of such a state presuppose the digital literacy of its citizens.

3.2 The Greek constitutional framework

Article 5A(2) of the Greek Constitution guarantees participation in the information society. This provision reflects an effort on the part of the revising legislator to embrace the digital age by facilitating access to electronically transmitted information.

Article 16 guarantees the freedom of art and science, stipulating in paragraph 1 that art and science, research and teaching shall be free, and that their development and promotion shall constitute a duty of the state. This provision may be interpreted to mean that the freedom of research is reinforced through the proper use of technology (e.g., research on incurable diseases may be supported by AI), teaching may be enhanced by AI tools, and art may be elevated by relevant applications, for instance, those that blend several artistic styles into a single work.

Article 16(2) stipulates that education constitutes a fundamental mission of the state. The constituent legislator confines itself to a general formulation, leaving it to the ordinary legislator to determine the means by which education is provided. It is therefore within the discretion of the ordinary legislator to establish the method of instruction, whether that involves digital tools and AI-assisted learning, the integration of technological means, or the maintenance of conventional educational models.

³ Case “Relating to certain aspects of the laws of languages in education in Belgium” v. Belgium (Merits), <https://hudoc.echr.coe.int/tur?i=001-57524>, last accessed: June 11, 2025.

⁴ European Court of Human Rights, Case of Hasan and Eylem Zengin v. Turkey, no. 1448/2004, Judgment of 9 October 2007

⁵ cf. Court of Justice of the European Union (CJEU), Maniero, C-457/17, Judgment of 15 November 2018, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:62017CJ0457>

At this point it should be noted that in a modern digital state, the right to education must necessarily encompass digital education. It would be unreasonable for public services to undergo digital transformation while individuals are not afforded the requisite education to navigate these services effectively. Such digitization, in the absence of corresponding educational measures, would be unconstitutional, as it would lead to a digital divide separating those who are able to acquire digital literacy through their own means and those who are not, thereby hindering or even obstructing their access to state services. Therefore, the right to digital education is inferred from Article 16(2), in conjunction with Article 5A, regarding access to the information society and the consequent avoidance of a “digital divide.” A lack of access to technology for educational purposes would run counter to the right to digital literacy. At the same time, the improper use of AI tools in education would also be unconstitutional, particularly when such use prevents certain population groups from accessing educational institutions that would otherwise support their intellectual, professional, economic, and social advancement.

Furthermore, the same paragraph 2 of Article 16 of the Constitution states that education shall aim at the moral, intellectual, professional, and physical development of the Greek people. From this provision, it may be inferred that the moral and intellectual cultivation of Greeks could also be achieved through the ethical use of technology. A prominent example of such ethical use of technology is the prohibition of plagiarism and the precise detection of textual similarity through AI tools. Technology may be deemed ethical when it is designed in a way that encourages its users to engage in moral behavior, rather than pushing them to commit criminal acts (as is the case with certain addictive online games).

Article 16(4) establishes the right to free public education. This obligation corresponds to the social right of pupils and students to receive education free of charge—a right that is closely aligned with the right to freedom of education. It is considered the oldest social right recognized under Greek constitutional law, having first been enshrined in the Constitution of 1864. At first glance, the provision of digital education may appear to be in tension with the State’s obligation to provide free public education, since not all pupils and students have access to modern digital learning tools. Social reality, however, largely dispels such concerns, as the majority of young people are users or owners of device suitable for education through AI-based tools. Complementary measures, such as providing educational devices to vulnerable groups and offering educational technology tools free of charge, further mitigate the risk of excluding socioeconomically disadvantaged populations. This is because the right to free public education does not only include the constitutionally mandated free provision of education by state educational institutions, but also the right of learners in need of assistance to receive financial support.

3.3 Regulation (EU) 2024/1689 on Artificial Intelligence (AI Act)

The significance of AI in the field of education is already highlighted in the Recitals of the Regulation. According to Recital 4,

“artificial intelligence is a fast-evolving family of technologies that can contribute to a wide array of economic, environmental and societal benefits across the entire spectrum of industry and social activities. By improving forecasting, optimizing operations and resource allocation, and personalizing digital solutions available to individuals and organizations, the use of AI can provide key competitive advantages to businesses and support socially and environmentally beneficial outcomes—for example in the areas of [...] education and training.”

According to Recital 56, “the deployment of AI systems in education is important to promote high-quality digital education and training and to allow all learners and teachers to acquire and share the necessary digital skills and competences, including media literacy, and critical thinking, to take an active part in the economy, society, and in democratic processes.” The Regulation classifies AI systems used in education or vocational training as high-risk. By way of example, it states that AI systems used for determining access or admission, for assigning persons to educational and vocational training institutions or programs at all levels, for evaluating learning outcomes of persons, for assessing the appropriate level of education for an individual and materially influencing the level of education and training that individuals will receive or will be able to access or for monitoring and detecting prohibited behavior of students during tests should be classified as high-risk AI systems, since they may determine the educational and professional course of a person’s life and therefore may affect that person’s ability to secure a livelihood. According to the Regulation, the improper design and use of systems may render them highly intrusive and violate the right to education and training, as well as the right not to be discriminated against. They may thus perpetuate historical attitudes of discrimination, for example against women, certain age groups, persons with disabilities or persons of certain racial or ethnic origins or sexual orientation.

Recital 96 underscores the importance of conducting an impact assessment when introducing AI systems into the field of education. The purpose of a fundamental rights impact assessment is to enable the implementing entity to identify specific risks to the rights of the individuals or groups of individuals likely to be affected, and to determine the measures to be taken in the event such risks materialize. The impact assessment must be conducted prior to the development of the high-risk AI system and should be updated whenever the implementing body considers that any of the relevant factors have changed. A critical issue arises with respect to who is responsible for assessing the risk—in other words, who determines what constitutes a risk. The impact assessment must identify the relevant processes of the implementing body in which the high-risk AI system will be used, in accordance with its intended purpose. It must also include a description of the timeframe and frequency of the system’s intended use, as well as the specific categories of natural persons and groups likely to be affected in each particular context of use. The assessment must also include the identification of specific risks of harm that are likely to have an impact on the fundamental rights of the individuals or groups concerned. When conducting the assessment, the implementing entity must take into account all information relevant to a proper assessment of the impact, including information provided by the provider of the high-risk AI system in the instructions for use. In light of the identified risks, implementing entities should determine

the measures to be taken in case such risks materialize. These may include, for example, governance arrangements within the specific context of use, such as provisions for human oversight in accordance with the instructions for use, or complaint handling and appeal procedures, as they could be instrumental in mitigating risks to fundamental rights in specific use cases. After carrying out the impact assessment, the implementing entity must inform the relevant market surveillance authority. Where appropriate, in order to collect relevant information necessary for carrying out the impact assessment, implementing entities of high-risk AI systems, in particular where the AI systems are used in the public sector, could ensure the involvement of relevant stakeholders. These include representatives of groups of people likely to be affected by the AI system, as well as independent experts, particularly during the impact assessment process and the design of mitigation measures to be adopted in case the risks materialize. The European Artificial Intelligence Office (AI Office) should develop a standardized questionnaire template in order to facilitate compliance and reduce the administrative burden on implementing entities.

The AI Regulation categorizes risks into four main categories. AI systems that present only limited risk will be subject to very light transparency obligations, while high-risk AI systems will require authorization and will be subject to a comprehensive set of requirements and obligations to gain access to the EU market. Certain AI systems, such as cognitive-behavioral manipulation and social scoring, will be prohibited by the EU, as their level of risk is deemed unacceptable. Under Article 5, the Regulation prohibits certain AI applications that pose a threat to individual rights. This reflects a consensus on the unacceptability of dangerous systems. In the field of education, emotion recognition is not permitted, except for medical or safety related reasons. This means that the emotional state of learners may not be monitored or analyzed during an educational process. For example, when learners are taking an exam, it is not permissible to record their emotions or draw inferences about them.

Under strict conditions, high-risk practices are permitted, as referred to in Article 6 et seq., such systems require an impact assessment. The Regulation sets out clear obligations for high-risk AI systems (due to the significant potential harm they may cause to health, safety, security, fundamental rights, the environment, democracy, and the rule of law). These systems must assess and mitigate risks, maintain usage logs, ensure transparency and accuracy, and provide for human oversight. Users will have the right to lodge complaints about the systems and receive explanations about decisions based on high-risk systems that affect their rights (Fanarioti and Karpouzis, 2025). The regulatory concept of high risk has its origins in the safety of the products concerned. High-risk practices in the field of education include, according to Article 6(c), the determination of access, admission or assignment of educational and vocational training at all levels, assessment of learning outcomes and the appropriate level of education, monitoring and detection of prohibited behavior of student during examinations. This does not mean that learners cannot be assessed and trained through technical means—but when such methods are used, appropriate safeguards must be in place to protect their rights.

According to Annex III, high-risk AI systems referred to in Article 6(3) include:

- AI systems intended to be used to determine access or admission or to assign natural persons to educational and vocational training institutions at all levels.
- AI systems intended to be used to evaluate learning outcomes, including when those outcomes are used to steer the learning process of natural persons in educational and vocational training institutions at all levels.
- AI systems intended to be used for the purpose of assessing the appropriate level of education that an individual will receive or will be able to access, in the context of or within educational and vocational training institutions at all levels.
- AI systems intended to be used for monitoring and detecting prohibited behavior of students during tests in the context of or within educational and vocational training institutions at all levels.

3.4 The proposal of the high-level expert group on AI

The Greek High-Level Expert Group proposes the development of a centralized AI education platform that will support teaching, learning and online collaboration, and host competitions in the field of AI. It recommends the provision of AI-related educational material through a centralized online platform. The platform will act as a common virtual space where educational material can be developed by AI specialist teams from academia and industry. Content creators will be invited to produce relevant material and will be remunerated based on its use. The aim is to promote the creation of a dynamic and sustainable ecosystem for AI education, where contribution is rewarded, and continuous improvement is encouraged. Educational material will be subject to a rigorous evaluation process to maintain a high level of quality and educational value. Educators will be able to select materials tailored to their teaching and audience, while individual learners will be able to pursue personalized studies to master new areas of knowledge and seek support to address any learning gaps. The same platform could also host competitions and hackathons (app development marathons), offering a virtual collaborative space. These activities could be organized around specific topics or challenges, encouraging participants to develop innovative solutions using AI applications. The platform would facilitate project submission and evaluation, provide access to tools and relevant datasets, and enable communication between teams and mentors. By hosting competitions and hackathons, the platform will encourage the creation of a community of students and researchers, enhance hands-on learning and inspire creativity and innovation in the field of AI and beyond. Finally, the same infrastructure can be used for vocational education and training, as well as for lifelong learning.

4 Case studies

While legal frameworks and human rights principles establish the necessary boundaries for the use of AI in education, their significance becomes clearer when examined in real-life situations. Case studies show how schools, universities, and learners experience AI applications in practice, showing both their promise and their pitfalls: they reveal how AI can improve administrative processes, support learning, and expand access, but also how it can raise concerns about fairness, surveillance, and the erosion of human judgment. The following examples, ranging from examinations and admissions to Generative AI in the classroom, illustrate how digital tools are already affecting educational practice and the everyday interactions between students and educators.

4.1 Conduct of examinations

The future of the examination process appears to be automated (Cerratto Pargman et al., 2023). It is anticipated that, in a few years, candidates for the national examinations will arrive at designated testing centers, sit for exams on dedicated computers, respond to standardized questions that do not require essay writing, and receive their grades upon completion of the testing session. Then, they will submit their preferences for academic institutions and subsequently receive a message indicating the school to which they have been admitted. At a later stage, examinations may be conducted remotely from the candidate's homes.⁶ In such cases, the examinee would be required to install specialized proctoring software on their personal computer or on a device provided by the examining authority to participate in the examination. The software will scan the exam environment, ensure that no third person is present at the examination site, perform a system scan of the examinee's computer, deactivate any suspicious software or communication tools and monitor the examinee's behavior throughout the process.

There are two risks involved in this process. First, the examinee's capacity for analytical writing is limited, as they have to answer multiple-choice questions. This may also restrict critical thinking, since it involves selecting from predefined answers. Secondly, there is a risk of excessive surveillance of the examinee. The first risk can be mitigated by asking questions that require analytical reasoning, even if it results in a slower grading process. Furthermore, software tools already exist that are capable of evaluating written responses. It could also be argued that critical thinking is not necessarily hindered when questions call for evaluative judgment; rather, such formats may increase the likelihood of examinee confusion or misinterpretation. The second risk could be addressed by taking technical and organizational measures to ensure the security of data processing.

⁶ This has already occurred during the examination process for the selection of senior executives in public administration

4.2 Evaluation of student applications

Admission to higher education institutions should require a comprehensive assessment of candidates. It is not sufficient to rely solely on entrance exam scores; additional factors, such as community involvement, individual skills, personality traits, and so on, should also be taken into account. Automated assessment of all these parameters can help minimize the influence of favoritism or nepotism and enhance the integrity of the admissions process. Nevertheless, it is not always easy for an algorithm to evaluate a candidate's soft skills, including qualities such as teamwork, honesty, and willingness to cooperate. The safest approach may be to rely on measurable criteria only, such as entrance exam scores and documented points earned through participation in athletic, musical, or other competitive activities. As for community service, it is proposed that a dedicated committee certify such contributions, issuing an official attestation that would be scored according to its duration and weighted appropriately in the admissions process.

4.3 (Pre-)assessment of faculty applications

It is envisaged that in the future the pre-assessment of candidates for faculty positions will be conducted in an automated manner. An algorithm will scan the application, verify whether the required formal qualifications have been met, for example, whether the candidate has completed mandatory military service, submitted their doctoral degree, or provided the required recognition for foreign qualifications, and so on. The number of publications will then be calculated and scored based on the impact or recognition of the scientific journal in which they are published. Such evaluation, however, is not always straightforward in the social sciences, where standardized metrics of journal significance are often lacking. Publications in peer-reviewed journals carry different weight compared to those in non-peer-reviewed ones. In Greece, moreover, there have been cases where journals claim to conduct blind peer review, yet the process is not genuinely anonymous. For these reasons, fully automated pre-screening of applications is not advisable. In any case, candidates retain the right to human intervention, as guaranteed under Article 22 of the GDPR.

4.4 Text generation through Generative AI

Generative AI has the potential to produce complex scientific answers to user-submitted questions (prompts), leading to applications that can provide automated responses to queries through the use of Large Language Models. These models are a type of AI algorithm that uses deep learning techniques and big data to produce text in a way that resembles human language. They are trained on massive data sets to predict likely text completions. The content and form of such responses are becoming increasingly significant

across multiple use cases related to education (Karpouzis, 2024).

4.5 Supporting the educational process

Specialized tools can assist the educational process: some of them generate engaging presentations, formulate possible questions with the desired degree of difficulty, produce summaries of course materials, evaluate students' answers, and focus on each learner individually. They offer personalized guidance, recommend revision in areas of weakness, carry out statistical evaluation of performance, identify gaps and deficiencies, check for text similarity, and relieve educators of administrative burdens. Such tools are valuable aids for those involved in education—provided they do not replace the teacher.

4.6 Examples of AI tools in education

4.6.1 Virtual education advisor

A widely used AI application in the field of education is a virtual education advisor: this tool can provide feedback on students' learning activities and practice questions and subsequently provide recommendations on which materials should be reviewed, much like a teacher or instructor would (Fitria, 2021). The system continuously learns and updates its information based on the needs and constraints faced by learners. It can also identify the underlying causes of student misunderstandings and suggest strategies to address learning difficulties.

In addition, the tool assists the learner during study sessions by helping to structure their thinking and by posing potential questions. In this way, the need for supplementary private tutoring may be reduced, as the learner no longer depends on an external instructor. Notwithstanding the above, it should not be viewed as a substitute for the educator, but rather as a support mechanism for both teachers and students.

4.6.2 Voice assistants

Voice assistants are among the most widely recognized and commonly used AI technologies in various fields, including education. Examples of well-known voice assistants include Google Assistant (Google), Siri (Apple), and Cortana (Microsoft); in an educational context, voice assistants allow learners to search for materials, reference questions, articles and books, simply by speaking or providing keywords.

4.6.3 Automated online assessment

AI is widely used for purposes of automated online assessment and grading of questions. The use of such features makes it easier for educators to prepare and administer tests easily and practically. Teachers no longer need to manually compose questions and correct answers. Instead, they simply select the type of questions, the level, number of items, difficulty, and

other parameters. Once the test is generated, the teacher can share the link with students who complete it online. This functionality enables the easy creation of assessment questionnaires (quizzes). Student results are instantly available in the teacher's account, with an overall score, a list of incorrect and correct answers, and a discussion feature. All of this is managed by a programmed AI system. AI technology can also support teachers by handling repetitive administrative tasks, such as lesson planning, exam grading, homework review, and more. By automating these processes, teachers gain more time to monitor student progress and focus on improving their instructional techniques.

This tool will operate autonomously based on programmed instructions and will be capable of learning from the user's or student's habits. Furthermore, the AI will provide recommendations for targeted material to be reviewed, as well as other suggestions based on the student's recorded performance.

4.6.4 Personalized learning

Personalized learning enables learners or users to receive services similar to those provided by personal assistants. The application of this technology is already quite common. AI allows learners to access tailored support by collecting data from their past learning activities and offering alternative learning paths based on individual needs. This approach enables each student to progress and develop at their own pace and according to their capacity to absorb content, in alignment with their interests and abilities. The AI will also provide content recommendations (Karpouzis et al., 2024), suggest to the user a timetable for study and various other important functions. Over time, the system learns how to optimize the learning process, making it more effective and efficient. By analyzing student data, AI can help educators and educational institutions identify each learner's pace and needs. Schools can then design study plans based on the students' strengths and weaknesses. What must be emphasized, however, is that the technology will function solely as a tool, allowing educators the time and space they need. In any case, the unique teacher-student relationship remains essential, especially when it comes to the emotional and ethical dimensions of learning, which directly affect students' feelings and psychological wellbeing.

4.6.5 Educational/serious games

Educational/serious games are designed for learning purposes, while still offering play and entertainment. They provide an educational or learning experience for the players. For example, Duolingo doesn't just teach English: it offers access to more than 30 foreign languages that children can learn, such as Mandarin, French, Italian, Spanish, Korean, Japanese, and others. Khan Academy Kids features thousands of interactive activities for toddlers, preschoolers and kindergartners. Within this all-in-one educational game, children can develop skills in reading, language, writing, math, social-emotional learning, problem-solving skills, and motor development. Quick Brain, on the other hand, sharpens the brain's processing speed for performing calculations.

4.6.6 Automatic text translation

Automatic text translation tools support those involved in the educational process by helping them understand the views of distinguished foreign-language representatives of science, literature, and the arts, by overcoming language barriers.

4.6.7 Virtual reality tools

Virtual reality tools can familiarize learners with foreign cultures and offer them an interactive experience of past historical eras (Marougkas et al., 2023). Learners may, for instance, engage in conversations with avatars of ancient philosophers and exchange ideas with them. By incorporating AI techniques, these platforms and their smart applications can provide personalized and immersive learning journeys, as well as automated and adaptive enhancements of the services offered—tailored to users' preferences, traits, and behavioral patterns. The user interacts with the environment and visualizes the information provided, resulting in enhanced understanding and assimilation of complex ideas and concepts. A key feature of virtual reality is the possibility of individualized guidance, with users being able to communicate directly with experts and receive real-time information adapted to their needs and interests.

4.6.8 Distance learning tools

During the pandemic, schools remained active thanks to the use of distance learning. These tools help overcome barriers to instruction during times of crises, such as pandemics, extreme weather, strikes, school closures, and so on. They enable learners to stay connected to the educational process. That said, they should not become a default or permanent substitute for in-person education, as there is a risk of weakening the social and relational bonds between educators and learners.

4.6.9 Text editing tools

Text editing tools assist with grammatical, syntactical and lexical correction, and vocabulary refinement. On the one hand, they help improve the quality of written output; on the other hand, overreliance on them may lead learners to neglect their own writing skills, knowing that an automatic corrector will always be available. One way to address this issue would be for the tool to identify the error but give the learner the opportunity to correct it manually before any automatic correction is applied.

5 Challenges and considerations

The incursion of AI into the field of education must not overlook the fundamental social function of education, i.e., the transmission of the capacity for critical engagement with knowledge and, more broadly, the teaching of critical thinking.⁷

⁷ National Committee on Bioethics and Technoethics, Opinion on the Applications of Artificial Intelligence in the Greek School, <https://bioethics.gr/announcements-26/nea-gnwmh-gia-tis-efarmoges-texnhths-nohmosynhs-sto-ellhniko-sxoleio-18-martioy-2025-3219>, last accessed: August 22, 2025.

Replacing this inherently human capacity risks leading to the heteronomy of learners and undermining the free development of their personality. Moreover, we must not lose sight of the fact that education is far more than the mere dissemination of information: it also aims at fostering social skills and shaping responsible and conscientious citizens, which is achieved through interaction and interpersonal engagement between students and educators. Therefore, any assessment of AI applications in education must begin with the premise that these tools are meant to be supportive in nature and not intended to replace the educator.

On one hand, AI tools can relieve educators of the burdensome administrative tasks, allowing them to focus more on student engagement and the cultivation of a positive learning environment. Presentations and learning become more engaging. Grading is conducted in an impartial and automated manner, without bias. Student progress is monitored in a systematic way. AI can help learners to develop critical thinking and computational skills, boosting their productivity and creativity, and enhancing their adaptability to technological changes.⁸ AI also offers significant opportunities in terms of providing educational resources for young people with disabilities and special needs. For example, AI-based solutions, such as real-time live captioning can assist individuals with hearing impairments, while audio description can improve access for those with low vision.⁹ Education in AI can also cultivate a deeper understanding of the ethical issues arising from its use, and of how to address them. Lastly, AI plays a vital role in facilitating broader access to knowledge through its many educational applications.

On the other hand, AI tools can lead to depersonalized teaching inaccurate assessments, and shortcomings in addressing the individual needs of each learner. Unregulated use of such tools can lead to fundamental errors that may mislead the learner from finding the truth. Mistakes can also occur during evaluation, as it is possible that multiple choice questions may not have been properly constructed. A virtual conversation with Socrates does not guarantee an accurate representation of his views. There is significant concern regarding the delivery of pre-designed, predetermined knowledge. Through poor—intentional or unintentional—training of the algorithm, Socrates may lead us toward a specific direction or a particular policy choice. Moreover, when students interact with digital devices, they generate digital traces. If not processed ethically, this kind of trace data (traces of digital usage and learning activity) can lead to an invasion of privacy. To this set of concerns, one can add the fear of excessive or unregulated processing of learners' personal data; the opacity or difficulty in explaining the outcomes produced by AI applications, especially in the case of machine learning systems; the challenge of incorporating ethical parameters and capabilities for logical reasoning and inference in the design of algorithms; the requirement that algorithms be fed with large volumes of scientific

⁸ Plan for Greece's Transition to the AI Era, https://foresight.gov.gr/wp-content/uploads/2024/11/Blueprint_GREECES_AI_TRANSFORMATION.pdf, last accessed: July 10, 2025.

⁹ European Union, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, <https://op.europa.eu/en/publication-detail/-/publication/d81a0d54-5348-11ed-92ed-01aa75ed71a1/language-en>, last accessed: July 10, 2025.

data; and the possibility that such data use may be restricted by intellectual property and industrial rights.

Whenever our society discovers a technological tool, a not unjustified concern arises that it will deprive learners of the opportunity to practice. When handheld calculators were introduced, there was a fear that students would lose their ability to solve mathematical problems. When search applications were created, concerns were raised that learners were being spoon-fed information. The same applied for tools providing bibliographic references. It should not escape our attention, however, that these tools facilitate education by allowing learners to engage in more complex issues. As with previous technologies, the question is not whether AI should be excluded, but how it can be integrated responsibly: the solution lies in the regulated and ethical use of such tools. This means that learners should not rely on using the tools without performing their own final review. As a result, the tool should assist and not replace the educator. Moreover, it is very useful that learners do not view the use of such tools as a magic genie that will solve all their problems. They should be assessed critically in an environment where the use of such tools is not allowed, be examined orally on the work they produce with the help of AI and be able to demonstrate that they have understood the educational content.

The goal is not to exclude AI tools from the field of education, but to ensure their proper and ethical use. The integration of AI tools into education should not be driven merely by their availability but based on their proven usefulness. This entails the following:

- Algorithms must be compatible with human dignity: AI systems should be developed in ways that respect the personal autonomy of those interacting with them. It is essential to take steps to prevent AI systems from exploiting, degrading, manipulating, instrumentalizing, or eroding human self-determination. This principle entails the exclusion of any applications that manipulate student behavior from the educational system. Practices involving the monitoring of student behavior, whether inside or outside the school environment; “social scoring” mechanisms; or the disclosure of either behavioral data or views expressed by students in class to third parties, when carried out through AI applications, violate the core of their personality rights.
- Algorithms must promote the wellbeing of those involved: their purpose is to support the educational process, not to replace the educator with an algorithm, and they should be recognized as tools that help and inspire people to improve their quality of life by utilizing their unique human capabilities in the context of education. They should also not be seen as a means of replacing educational human labor for the sole purpose of reducing costs.
- Algorithms must be transparent: we should be able understand how they work, especially when processing (e.g., evaluating or summarizing) texts, which may alter their original meaning (cf. the Section on XAI). Algorithms, data, and AI-based decision-making procedures must be sufficiently accessible to relevant stakeholders so that the functioning of AI systems is understandable, explainable, reliable, justified, and accountable.

- Algorithms must respect the privacy of learners: recording of student progress, performance fluctuations, periods of fatigue, or emotional stress must be handled in a manner that respects the student’s rights and does not lead to negative consequences in their future personal and professional development.¹⁰ Educational institutions are required to ensure that all data they process are stored confidentially and securely and must implement appropriate policies and procedures for the protection and ethical use of all personal data (European Commission and Culture, 2022).
- Algorithms must promote pluralism: this means that they must provide learners with pluralistic education, be inclusive, and encompass a wide range of disciplines, from mathematics and natural sciences to the humanities and social sciences (Karpouzis, 2024). They must not promote cultural and linguistic monoculture; instead, they should foster diversity and social cohesion and aim at inclusion.
- Algorithms must be monitored and supervised, in order to weigh the risks and benefits to stakeholders, and promote the required values. Special attention should be paid to the choice of the appropriate form of oversight (Panagopoulou, 2024), including human supervision.
- A distinction of particular importance must be made between education in AI and education about AI; education in AI refers to the integration of AI into curricula as part of digital and algorithmic literacy. What is necessary, however, is a shift of focus toward education about AI: that is, familiarizing students with the ethical, social, and legal questions raised by AI, with the ultimate goal of preparing them for a future of responsible and constructive interaction with AI systems that are increasingly embedded in everyday life.
- Algorithms should not have as their primary mission the mere transmission of information, but rather the promotion of critical thinking.
- Algorithms must ensure equal access to AI-based applications for all learners.
- Algorithms must be used in a manner that complements, rather than replaces, direct teacher-led instruction. AI applications in education must not replace the student–teacher relationship, nor the interpersonal bonds that form the fabric of the educational community. The school group, the classroom, and the wider school community must remain the primary environments for shaping students’ social identity and the development of their social skills.

¹⁰ S. and Marper v. United Kingdom of the European Court of Human Rights (ECtHR, Applications nos. 30562/2004 and 30566/2004, Judgment of 4 December 2008) also moves in this direction. According to the Court, the retention of data concerning individuals who have not been convicted may be particularly harmful in the case of minors, due to their specific situation and the importance of their development and integration into society. The Court held that special attention must be paid to the protection of minors from potential harm arising from the retention of their personal data by the authorities following acquittal or non-conviction.

6 Integrating explainable AI in education

The aforementioned examples of AI tools, along with their associated challenges and considerations, provide evidence that AI is revolutionizing the education industry worldwide, gradually reshaping how educators teach and students learn. The use of AI in collaborative teacher–student learning, intelligent tutoring systems, automated assessment, personalized learning, and real-time student feedback is poised to shift the paradigm in teaching and learning (Kamalov et al., 2023). AI technologies personalize teaching experiences, enable teachers to focus on more strategic aspects of teaching, improve students' learning experiences, enhance engagement and performance, and ultimately facilitate both general and higher education (Jian, 2023), thereby influencing students' academic development (Vieriu and Petrea, 2025). As is readily seen, AI applications are transforming the educational landscape; thus, it is of crucial importance to ensure the sustainable development and deployment of AI-driven technologies at schools and universities (Kamalov et al., 2023). However, while AI is particularly useful in heterogeneous learning environments, enabling scalable and adaptive solutions for teaching and learning, a systematic review of AI applications in higher education, including educational institutes with learners coming from varied backgrounds, abilities, and personal attributes, revealed limited discussion on its practical integration as well as its ethical, legal, and technological challenges, inadequate alignment with general pedagogical theories, and a clear need for a more in-depth investigation into the ethical and educational implications of its use (Zawacki-Richter et al., 2019).

6.1 Equitable, resilient and transparent digital educational AI ecosystem

The Horizon Report on higher education,¹¹ education's longest-running exploration of emerging technology trends, forecasted that AI applications related to teaching and learning is projected to grow significantly, expected to support teaching, learning, and creative inquiry, while also raising ethical concerns and risks associated with its development. This report not only profiled emerging educational technology for higher education, but also identified key trends and significant challenges, serving as a reference and technology planning guide for educators, higher education leaders, administrators, policymakers, and technologists. The important developments in educational technology for higher education were classified based on the time-to-adoption horizon, one year or less for “mobile learning” and “analytics technologies,” two to three years for “mixed reality” and “artificial Intelligence,” and four to five years for “blockchain” and “virtual assistants.” The key trends accelerating technology adoption in higher education were sorted along a time continuum. Long-term trends included “rethinking how institutions work”

and “modularized and disaggregated degrees,” mid-term trends included “advancing cultures of innovation” and “growing focus on measuring learning,” and short-term trends included “redesigning learning spaces” and “blended learning designs.” Moreover, the significant challenges impeding technology adoption in higher education were classified into three categories. Solvable challenges, those that we understand and know how to solve, included “improving digital fluency” and “increasing demand for digital learning experience and instructional design expertise,” difficult challenges, those that we understand but for which solutions are elusive, included “the evolving roles of faculty with EdTech strategies” and “achievement gap,” and wicked challenges, those that are complex to even define, much less address, included “advancing digital equity” and “rethinking the practice of teaching.” Furthermore, UNESCO (Unesco, 2015) placed great emphasis on the need to ensure that AI warrants educational equity and inclusion, as a means to achieve the Education 2030 Agenda, rather than reinforcing existing digital divides, especially across different regions and socioeconomic groups. In response to the vision of “AI for all,” panels of expert practitioners and stakeholders have established regulatory frameworks for the ethical governance of AI in education (e.g., Australia's AI Ethics Framework¹²) and AI policy education frameworks for university teaching and learning (Chan, 2023), in order to foster a comprehensive understanding of the multifaceted implications of AI integration in academic environments. Three key dimensions emerged from these policy reports: *pedagogical*, focusing on utilizing AI to enhance teaching and learning outcomes; *governance*, for handling privacy, security, and accountability concerns; and *operational*, addressing infrastructure and training issues.

6.2 The role of XAI in education

As can be seen, there is a strong demand, now more than ever, for fostering a digital pedagogical environment that is ethical, inclusive, dynamic, resilient, and transparent. Addressing the “black-box” problem is one of the largest issues in AI, particularly in learning settings. In most machine learning and AI systems, decisions and choices are made following a black-box approach, in manners which are not easy to comprehend for humans, and thus it becomes difficult for instructors and students. This alone is enough to decrease the level of trust on these systems, which is needed for their extensive usage. Computer algorithms and adaptive learning systems, for instance, will at times offer feedback or scores instead of providing a transparent and intelligible justification, thereby lacking reasonableness behind their decisions. Although some interpretability techniques,¹³ such as Local Interpretable Model-agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP), have been developed, they often fail to provide the straightforward, actionable insights that

11 EDUCAUSE Horizon Report: 2019 Higher Education Edition, <https://library.educause.edu/-/media/files/library/2019/4/2019horizonreport.pdf>, last accessed: August 22, 2025.

12 Artificial Intelligence: Australia's Ethics Framework, <https://www.csiro.au/en/research/technology-space/ai/AI-Ethics-Framework>, last accessed: August 22, 2025.

13 These are statistical techniques that attempt to “open the black box” of AI systems by explaining why they made certain predictions

educators need (Hooshyar and Yang, 2024). Solution to this issue requires AI systems to be designed based on human-centered principles with a focus on explainability in an effort to build trust and enable all users to become proficient in using such technology. Due to the classification of education as a “high-risk” sector in regulatory frameworks like the EU AI Act, accurate and reliable interpretability in education-specific AI applications is more crucial given the vast influence educational AI has on the pace of learning and students’ academic development (Karpouzis, 2023).

Moreover, from the moment that AI technology has been integrated into Learning Management Systems (LMS) and Educational Data Mining (EDM) tools, either for supporting the learning process and/or for evaluating students’ performance, there has been a growing demand for explainability (Rachha and Seyam, 2023), with Explainable AI (XAI) in education (Khosravi et al., 2022) playing a leading role for enhancing transparency of AI-based decision-making. Describing and reasoning on how decisions and recommendations are made, address effectively the “black box” nature of traditional AI providing the needed transparency in education-oriented decision making which is critical in learning environments, where learners and educators need to know the reasoning of the AI to ensure fairness, effectiveness, and inclusivity. For example, in adaptive learning platforms, XAI can describe why a given resource was suggested to a learner or how considerations such as past knowledge and/or engagement influenced a proposed learning path. This explainability not only allows teachers to verify AI’s recommendations, but also to ensure alignment with learning standards and warrant fairness in students’ evaluations.

There are, however, several challenges since XAI tools, such as the previously mentioned LIME and SHAP, require technical skills, perhaps not so readily available to various existing educational environments. Moreover, the explanations generated by such AI means can be difficult to contextualize without considering the broader educational ecosystem. Although ongoing research emphasizes that XAI in education shares similar status and characteristics as its application in other fields, it also imposes some special demands for tailored solutions. Therefore, it is important to trace the evolution of XAI techniques from the initial stages to the latest developments, and to address the most critical challenges in applying these methods in education. Doing so will allow delivering efficient explanations to learners and rendering AI models transparent and explainable to educators, so that the results of AI can be utilized to guide instructions in the proper manner. Moreover, the complexity and context-sensitivity of educational decision-making are two main characteristics that XAI-driven technology should integrate in the design of its educational AI tools in a way that those complement human judgment, rather than replacing it, and ensuring that AI serves as a complementary tool in learning and assessment processes in educational settings.

Although the aforementioned concerns are widely recognized, there is still no worldwide consensus about the management strategies required to ensure that AI tools are not only designed but also applied responsibly across the diverse digital educational ecosystem. Although the European Union’s AI Act, the first-ever legal framework on AI, addresses the risks of AI and positions

Europe to play a leading role globally, it still remains a high-level policy that lacks explicit guidelines for educators and AI developers. In addition, its emphasis on ethics inevitably leads to overlooking some of the practical issues of implementing these principles within various learning environments. UNESCO’s educational strategy and competency frameworks for AIXXX, further highlight the importance of educators and learners both being informed for the opportunities and threats of AI, so as to be equipped with the necessary competencies to use these technologies responsibly. Addressing these challenges requires robust policies and practical support to deliver fair, open, and responsible XAI-driven education. It is therefore clear that a comprehensive framework laying down harmonized rules on XAI is needed to foster its development, use, and uptake.

7 Policy recommendations

Toward this direction, in October 2024, the European Digital Education Hub of the European Commission held a workshop in which 31 AI experts participated to investigate the educational implications of XAI. Their main role was to evaluate explainability methods and refine AI systems’ integration into educational contexts based on practical examples. The regulatory framework of the AI Act, the crucial role of ethics, and the necessity of AI literacy were discussed, and recommendations were then made, such as developing an AI literacy framework and adapting the EU’s SELFIE tool¹⁴ to measure AI transparency. This workshop serves as a leading example of XAI-oriented initiatives, alongside other global efforts, from which several policy recommendations have been drawn for integrating XAI in education; some of them are shown in Tables 1–3, grouped with respect to priority, implementation pathway and measurable Key Performance Indicators (KPIs). In these tables, short-term (0–12 months) measures rely on instruments already within ministerial or government discretion (AI-literacy programs, accountability guidance, standardized explainability scores embedded in procurement), while medium-term steps build durable capacity through inclusive-access schemes, multi-stakeholder co-creation, targeted R&D, and sector-specific ethics review. Long-term monitoring then embeds governance in routine inspection and public reporting. For each item we identify the minimum viable steps and measurable indicators, so policymakers and educators have a clear pathway rather than general aspirations.

However, these policy recommendations risk remaining aspirational rather than becoming reality lacking adequate support and resources, and require changes in pedagogy and educational infrastructure to be done. XAI has the potential to advance the education industry, but this can only be achieved to its full potential if significant efforts are undertaken to overcome legal, ethical, technical, and practical issues of its implementation in education.

¹⁴ SELFIE Toolkit for users, <https://education.ec.europa.eu/selfie-for-teachers/toolkit>, last accessed: August 22, 2025.

TABLE 1 Short term (0–12 months) recommendations.

Recommendation	Pathway (Minimum steps)	KPIs
Implement AI literacy in the education system at all levels so that learners and educators both have the knowledge and skills needed to critically engage with AI tools.	Publish teacher/student AI-literacy outcomes and a 6–8 h micro-credential; provide template lesson activities; align with existing digital-skills curricula.	PD completion rates; pre/post self-efficacy; % schools embedding at least one AI-literacy task/term.
Clarify accountability guidelines to facilitate educators and students comprehending their rights and responsibilities when using AI systems.	Define roles (provider, deployer, user), human-oversight expectations, complaint/appeal routes, record-keeping; add Fundamental Rights Impact Assessment (FRIA)/Data Protection Impact Assessment (DPIA) triggers for high-risk uses.	% institutions adopting policy; time-to-resolution for complaints; audit pass rate.
Standardize explainability by creating an “explainability score” for AI tools that would guide educators in selecting transparent and comprehensible AI systems.	Publish a rubric (inputs, model cards, user-facing rationale quality, error handling); require vendors to self-score; pilot in 3–5 typical tools (assessment, tutoring, analytics).	% procurements using the rubric; inter-rater reliability across pilots; educator satisfaction with explanations.
Design AI with ethical, human-centered values	Include dignity, autonomy, transparency, privacy, pluralism criteria; require risk/benefit trade-offs and mitigation.	% tools meeting baseline; # issues flagged in design reviews.

TABLE 2 Medium term (12–36 months) recommendations.

Recommendation	Pathway (Minimum steps)	KPIs
Support inclusive access to AI technologies, particularly to marginalized communities and institutions with limited resources by means such as subsidies, grants, or partnerships to bring AI tools and infrastructure within an affordable reach.	Identify high-need schools; negotiate educational licensing of technologies; publish equity dashboard.	Coverage in priority schools; device-to-student ratios; usage parity across socio-economic status groups.
Encourage multi-stakeholder collaboration to make sure that AI tools are co-created to meet the varied educational requirements.	Convene teachers, students, parents, SEND experts, DPAs to co-design; run 2–3 pilots per year with public reports.	Pilots completed; adoption of co-designed features; stakeholder satisfaction.
Fund research and development to warrant continuous innovation and the integration of AI systems in education.	Prioritize explainability, accessibility, and teacher-in-the-loop designs; require open evaluations and classroom trials.	Funded projects; open datasets/toolkits released; classroom impact reports.
Establish ethics review committees that will assess AI instruments in terms of meeting educational principles as well as ethical standards.	Scope reviews for pilot deployments; templates for proportional review; link to FRIA/DPIA.	Reviews completed; median review time; % deployments with oversight.

TABLE 3 Long term (36–60 months) recommendations.

Recommendation	Pathway (Minimum steps)	KPIs
Ongoing monitoring and assessment to make sure that AI systems are working as intended and responding to shifting educational needs and/or societal expectations.	Add AI-governance items to inspections (oversight, logs, transparency to learners), aggregate incident data nationally, publish annual report.	Compliance trend lines; incident recurrence; improvement between inspection cycles.

8 Conclusions

This is a legal-doctrinal and policy analysis rather than an empirical evaluation. Our emphasis on EU and Greek sources means that the argument is there, but jurisdictional details beyond these settings are discussed only selectively. We rely primarily on statutes, constitutional provisions, advisory opinions, and secondary literature; as such, we do not claim comprehensive coverage of case law across Member States or a systematic review of empirical effectiveness for specific AI tools. The case vignettes and examples (e.g., examinations, admissions, use of LMS/XAI) are indicative, not representative, and their legal implications may vary with institutional policies and local data-protection practice. In addition, the regulatory environment is moving quickly, notably around the AI Act’s implementation acts and guidance, so some details may date after our last source cut-off; future work should track delegated acts, national implementing measures, and test the proposed governance measures—AI literacy, explainability scoring, impact-assessment

routes—in real school and university settings. Finally, our review did not systematically assess ToM/metacognition claims in AI (e.g., Pergantis et al., 2025; Cuzzolin et al., 2020) and, more specifically, LLMs (Strachan et al., 2024); future work should test such methods in prospective, real-world studies (with preregistered outcomes) before treating them as safety or efficacy enhancements.

Overall, the use of AI in the field of education is recommended following strategic planning and careful assessment of the risks it entails. AI tools must serve the fundamental principles of protecting human dignity, privacy, transparency, promoting wellbeing, and pluralism. At the same time, they must be subject to oversight by supervisory authorities. Regarding their introduction into the educational process, it is first proposed that such tools should support, rather than replace the educator. Second, it is recommended that they be used as a complement to conventional education and not as a substitute for it (Panagopoulou et al., 2023). Traditional education, which is grounded in dialogue between teacher and learner, must not be abandoned, but rather

enriched with new educational methods. Particular emphasis must also be placed on the ethical use of technology within the educational sphere. And, lastly, AI must not divert us from the fundamental aim of education: the formation of responsible and conscientious citizens. In this framework, the legal and ethical integration of AI in education will determine whether it serves as a tool for empowerment or becomes a source of new inequalities.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

FP: Writing – review & editing, Conceptualization, Investigation, Methodology, Writing – original draft. CP: Writing – review & editing, Investigation, Writing – original draft. KK: Writing – original draft, Investigation, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

References

- Benzies, K. M., Premji, S., Hayden, K. A., and Serrett, K. (2006). State-of-the-evidence reviews: advantages and challenges of including grey literature. *Worldviews Evid. Based Nurs.* 3, 55–61. doi: 10.1111/j.1741-6787.2006.00051.x
- Cerratto Pargman, T., Lindberg, Y., and Buch, A. (2023). Automation is coming! Exploring future (s)-oriented methods in education. *Postdig. Sci. Educ.* 5, 171–194. doi: 10.1007/s42438-022-00349-6
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *Int. J. Educ. Technol. High. Educ.* 20:38. doi: 10.1186/s41239-023-00408-3
- Cukurova, M., and Miao, F. (2024). *AI Competency Framework for Teachers*. Geneva: UNESCO Publishing.
- Cuzzolin, F., Morelli, A., Cirstea, B., and Sahakian, B. J. (2020). Knowing me, knowing you: theory of mind in AI. *Psychol. Med.* 50, 1057–1061. doi: 10.1017/S0033291720000835
- Delors, J. (1996). *Learning: The Treasure Within*. Geneva: UNESCO.
- European Commission, Directorate-General for Education, Y. S., and Culture (2022). *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*. Publications Office of the European Union.
- Fanarioti, A. K., and Karpouzis, K. (2025). Artificial intelligence and the future of mental health in a digitally transformed world. *Computers* 14:259. doi: 10.3390/computers14070259
- Fitria, T. N. (2021). “Artificial intelligence (AI) in education: using ai tools for teaching and learning process, in *Prosiding Seminar Nasional Call for Paper STIE AAS*, 134–147.
- George, B., and Wooden, O. (2023). Managing the strategic transformation of higher education through artificial intelligence. *Admin. Sci.* 13:196. doi: 10.3390/admsci13090196
- Harris, D. J., O’boyle, M., Bates, E., and Buckley, C. (2023). *Law of the European Convention on Human Rights*. Oxford: Oxford University Press. doi: 10.1093/he/9780198862000.001.0001
- Holstein, K., and Doroudi, S. (2021). Equity and artificial intelligence in education: will “aied” amplify or alleviate inequities in education? *arXiv preprint arXiv:2104.12920*.
- Hooshyar, D., and Yang, Y. (2024). Problems with shap and lime in interpretable AI for education: a comparative study of post-hoc explanations and neural-symbolic rule extraction. *IEEE Access* 12, 137472–137490. doi: 10.1109/ACCESS.2024.3463948
- Jian, M. (2023). Personalized learning through AI. *Adv. Eng. Innov.* 5, 16–19. doi: 10.54254/2977-3903/5/2023039
- Kamalov, F., Santandreu Calonge, D., and Gurrib, I. (2023). New era of artificial intelligence in education: towards a sustainable multifaceted revolution. *Sustainability* 15:12451. doi: 10.3390/su151612451
- Karpouzis, K. (2023). “Explainable ai for intelligent tutoring systems, in *International Conference on Frontiers of Artificial Intelligence, Ethics, and Multidisciplinary Applications* (Springer), 59–70. doi: 10.1007/978-981-99-9836-4_6
- Karpouzis, K. (2024). Plato’s shadows in the digital cave: controlling cultural bias in generative ai. *Electronics* 13:1457. doi: 10.3390/electronics13081457
- Karpouzis, K., Pantazatos, D., Taouki, J., and Meli, K. (2024). “Tailoring education with genai: a new horizon in lesson planning, in *2024 IEEE Global Engineering Education Conference (EDUCON)* (IEEE), 1–10. doi: 10.1109/EDUCON60312.2024.10578690
- Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., et al. (2022). Explainable artificial intelligence in education. *Comput. Educ.* 3:100074. doi: 10.1016/j.caeai.2022.100074
- Margaritis, M. (2018). *The European Convention on Human Rights (ECHR) and Protocols Nos. 1, 6, 7, and 13: A Concise Interpretation*. Athens, Greece: P. N. Sakoulas.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Marougkas, A., Troussas, C., Krouska, A., and Sgouropoulou, C. (2023). Virtual reality in education: a review of learning theories, approaches and methodologies for the last decade. *Electronics* 12:2832. doi: 10.3390/electronics12132832
- O'Neil, C. (2017). Weapons of math destruction: How big data increases inequality and threatens democracy. *Crown* 44, 97–98. doi: 10.1177/0256090919853933
- Panagopoulou, F. (2024). Artificial intelligence and independent authorities. *J. Public Admin.* 6, 34–38. doi: 10.22259/2642-8318.0601004
- Panagopoulou, F., Parpoula, C., and Karpouzis, K. (2023). Legal and ethical considerations regarding the use of chatgpt in education. *arXiv preprint arXiv:2306.10037*.
- Pergantis, P., Bamicha, V., Skianis, C., and Drigas, A. (2025). AI chatbots and cognitive control: enhancing executive functions through chatbot interactions: a systematic review. *Brain Sci.* 15:47. doi: 10.3390/brainsci15010047
- Rachha, A., and Seyam, M. (2023). Explainable AI in education: current trends, challenges, and opportunities. *SoutheastCon 2023*, 232–239. doi: 10.1109/SoutheastCon51012.2023.10115140
- Selwyn, N. (2019). *Should Robots Replace Teachers? AI and the Future of Education*. New York: John Wiley Sons.
- Strachan, J. W., Albergo, D., Borghini, G., Pansardi, O., Scaliti, E., Gupta, S., et al. (2024). Testing theory of mind in large language models and humans. *Nat. Hum. Behav.* 8, 1285–1295. doi: 10.1038/s41562-024-01882-z
- Tamamidis, A. (2021). "Interpretation of article 2 of the first additional protocol to the ECHR, in *Article-by-Article Interpretation of the ECHR*, eds. X. C. Ioannis Sarvas, and C. Anthopoulos (Athens, Thessaloniki: Sakkoulas Publications), 1094.
- UN (1948). *Universal Declaration of Human Rights*. United Nations.
- UNESCO (2016). "UNESCO global action programme on education for sustainable development: information folder," in *UNESDOC Digital Library (ED/IPS/ESG/2017/02)* (UNESCO). Available online at: <https://unesdoc.unesco.org/ark:/48223/pf00000246270>
- UNG Assembly (1989). *Convention on the Rights of the Child*. United Nations, 1–23.
- Varsik, S., and Vosberg, L. (2024). *The potential impact of artificial intelligence on equity and inclusion in education*. Technical report, OECD Publishing.
- Vierdag, E. W. (1978). The legal nature of the rights granted by the international covenant on economic, social and cultural rights. *Netherlands Yearb. Int. Law* 9, 69–105. doi: 10.1017/S0167676800003780
- Vieru, A. M., and Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. *Educ. Sci.* 15:343. doi: 10.3390/educsci15030343
- Zawacki-Richter, O., Marin, V. I., Bond, M., and Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators? *Int. J. Educ. Technol. High. Educ.* 16, 1–27. doi: 10.1186/s41239-019-0171-0
- Zuboff, S. (2023). "The age of surveillance capitalism, in *Social Theory Re-Wired* (Routledge), 203–213. doi: 10.4324/9781003320609-27