

An initiative of the



European
Commission



AI and Ethics, Human Rights, Law and Educational Data

Briefing report No. 6

**by the European Digital Education Hub's squad on artificial
intelligence in education**

**EUROPEAN
DIGITAL
EDUCATION
HUB**

Content

| | |
|---|-----------|
| Introduction | 3 |
| Ethics of AI | 4 |
| Examples of National or Institutional Regulation on AI | 7 |
| Guidelines on Applying AI Specifically to Education | 8 |
| Data Protection | 11 |
| Cases Regarding Violation of Privacy and Data Protection in Education | 12 |
| Education and Ethics of AI | 14 |
| Recommendations by the Squad | 16 |





Introduction

In March 2023, more than a thousand technology leaders and researchers have [called for a halt](#) in the development of powerful artificial intelligence (AI) tools, pointing out the risks that the current race to develop more powerful AI could pose to society, especially considering the huge number of unknowns with these technologies and the lack of regulation (Future of life institute, 2023).

Following this letter, UNESCO has called for the immediate implementation of the [Recommendation on the Ethics of AI](#) that was unanimously adopted by its member states in November 2021. In [this call](#), the organisation raises its concerns about “many of the ethical issues raised by these innovations, in particular discrimination and stereotyping, including the issue of gender inequality, but also the fight against disinformation, the right to privacy, the protection of personal data, and human and environmental rights”. UNESCO also states that industry self-regulation is insufficient to avoid ethical harms and that AI developments should abide by the rule of law, avoiding harm, and that mechanisms should exist to ensure accountability and redressal (UNESCO, 2023).

The issues linked to ethics of AI, the right to privacy, data protection, gender inequality or human rights, are also present in the education sector, where the population is often more vulnerable, notably due to a young age and a lack of understanding. Therefore, it is highly important to put in place and implement legal safeguards and technical norms for the ethical use of AI in education, to ensure that its use does not violate rights of students, teachers and other people in the educational sphere. If this duty must be mainly the responsibility of the states, other actors, including schools, teachers as well as tech companies have an important role to play. Ensuring that students are aware of these issues is also important for them to understand how AI systems work and what their risks are.





Ethics of AI

Adopted in 2021, the [UNESCO Recommendation on the Ethics of Artificial Intelligence](#) is a non-binding international instrument, the first-ever global instrument to deal with the topic of ethics of AI. The Recommendation highlights four core values that should lay the foundations for AI systems and ten core principles that lay out a human-rights centred approach to the ethics of AI. The core values are:

- Respect, protection and promotion of human rights and fundamental freedoms and human dignity;
- Living in peaceful just, and interconnected societies;
- Ensuring diversity and inclusiveness;
- Environment and ecosystem flourishing.

The core principles are:

- Proportionality and doing no harm;
- Safety and security;
- Right to privacy and data protection;
- Multi-stakeholder and adaptive governance and collaboration;
- Responsibility and accountability;
- Transparency and explainability;
- Human oversight and determination;
- Sustainability;
- Awareness and literacy;
- Fairness and non-discrimination.

The Recommendation highlights that AI raises new types of ethical issues, including their impact on education and that new ethical challenges are created because of the potential of AI algorithms to exacerbate already existing biases and discrimination (paragraph 2c). It further states that it pays specific attention to education, “because living in digitalizing societies requires new educational practices, ethical reflection, critical thinking, responsible design practices and new skills, given the implications for the labour market, employability and civic participation” (paragraph 3a).

The Recommendation also gives concrete policy recommendations, including for education and research (policy area 8). It notably recommends states to provide adequate AI literacy education, to encourage research initiatives on the responsible and ethical use of AI technologies in teaching, teacher training and e-learning, to promote the leadership of girls and women, diverse ethnicities and cultures, persons with disabilities and vulnerable people, to develop AI ethics curricula or to ensure a critical



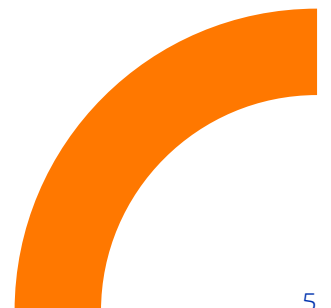


evaluation of AI research, and proper monitoring of potential misuses or adverse effects (UNESCO, 2022).

The draft regulation of the **European Parliament and of the Council**, laying down [harmonised rules on artificial intelligence \(AI Act\)](#) and currently under discussion, highlights the importance of deploying AI systems to modernise education systems, enhance educational quality both offline and online, and increase access to digital education for a wider audience. However, the use of AI systems in education, particularly for decisions regarding admissions, evaluations, and determining appropriate levels of education, carries ethical implications. These AI systems should be classified as high-risk due to their potential to shape an individual's educational and professional trajectory, impacting their ability to secure their livelihood. Improperly designed and used AI systems can be intrusive, violating the right to education, perpetuating discrimination, and reinforcing historical patterns of bias against certain groups, such as women, specific age groups, individuals with disabilities, or those of certain racial, ethnic groups, or sexual orientations.

AI systems identified as high-risk include AI technology used in critical infrastructures, and some cases in educational or vocational training that may determine the access to education and professional course of someone's life. Those are the systems that could put the life and health of citizens at risk ([EC, 2022](#)).

"AI systems used in education or vocational training, notably for determining access or materially influence decisions on admission or assigning persons to educational and vocational training institutions or to evaluate persons on tests as part of or as a precondition for their education or to assess the appropriate level of education for an individual and materially influence the level of education and training that individuals will receive or be able to access or to monitor and detect prohibited behaviour 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 affect their ability to secure their livelihood. When improperly designed and used, such systems can be particularly intrusive and may violate the right to education and training as well as the right not to be discriminated against and perpetuate





historical patterns of discrimination, for example against women, certain age groups, persons with disabilities, or persons of certain racial or ethnic origins or sexual orientation” ([AI Act proposal, May 2023, p. 113](#)).

Previously, the [European Parliament resolution on artificial intelligence in education, culture and the audiovisual sector](#) (2021), already called to include education in the regulatory framework for high-risk AI systems, “given the particularly sensitive nature of data on pupils, students and other learners” (European Parliament, 2021). In its general observations, this resolution also highlighted that the use of AI in education raises concerns about the ethical use of data, learners’ rights, data access and protection of personal data, hence putting at risk fundamental rights, including through the risk of creation of stereotyped models of learners’ profiles and behaviour that could lead to discrimination or doing harm by scaling up bad pedagogical practices (§ AD). The part dedicated to education, while recognising that AI can offer a wide range of great possibilities and opportunities in education, also points out the various issues that can arise and expresses notably the need for strengthening

digital skills, the fundamental role of teachers and the necessity to direct more public money towards AI research universities. It also states that there are specific risks to use AI automated recognition applications and that the European Commission should ban automated biometric identification such as facial recognition for educational purpose, unless its use is allowed by the law (§45).

It is interesting to note that the **Consultative Committee of the Convention 108 on data protection, issued guidelines on facial recognition** in 2021, ([Consultative Committee Convention 108, 2021](#)) in which it did not recommend to ban facial recognition in education but affect recognition. It stated that: “affect recognition can also be carried out with facial recognition technologies to arguably detect personality traits, inner feelings, mental health or workers’ engagement from face images. Linking recognition of affect, for instance, to hiring of staff, access to insurance, to education may pose risks of great concern, both at the individual and societal levels and should be prohibited” (1.1).





Examples of National or Institutional Regulation on AI

The Cyberspace Administration of **China** (CAC) [has released a draft](#) regulation titled “[Regulations on Governing the Service of Generative AI](#)” for consultation. The regulation aims to govern the use of Generative AI (GAI) services within the People’s Republic of China. It encourages innovation, trustable AI, and the use of safe software, tools, computing, and data sources. Discrimination based on race, ethnicity, religion, nationality, gender, age, or profession is prohibited throughout the AI design process. Intellectual property rights and business ethics must be respected, and unfair competition is not allowed.

GAI providers are expected to adhere to core values, avoid the unauthorised use of personal information and commercial confidentiality information, avoid unlawful content, and refrain from generating disinformation or content that may cause social or economic disruptions.

Providers are responsible for ensuring the authenticity, accuracy, objectivity, and diversity of the data. Transparency is emphasised, with providers required to disclose information about data sources, labelling, algorithms used, and user complaints mechanisms should be in place. Users should be provided with guidance on responsible AI usage and prevention of harm to others by providers.


Before offering GAI services to the public, providers must undergo security assessments and register with the Cyberspace Administration of China. Non-compliance with the regulation can result in penalties, including fines and service suspension.

On a different continent, the authorities in **Brazil** are in the process of drafting a new [framework for regulating the ethical and responsible use of Artificial Intelligence \(AI\) systems](#). This new bill, which will result in the replacing of three previous bills, comes about after a lengthy consultative process. The bill consists of eight separate chapters dealing in depth with the following points:

- Establishing national norms for ethical and responsible use of AI systems;
- Protecting individual rights;
- Risk based approach to AI regulation;
- Governance and algorithmic impact assessments of AI systems;
- Civil liability for damages caused by AI systems;
- Regulation and oversight of AI law.

Its primary aim is to grant individuals significant rights and place specific obligations on companies that develop or use AI technology (AI suppliers or operators). To achieve this, the bill establishes the creation of a new regulatory body to enforce the law and takes a risk-based approach by categorising AI





systems. It also introduces a protective system of civil liability for providers or operators of AI systems, along with a reporting obligation for significant security incidents.

Ethics is not only about defining principles top-down, but also about the discourse of the people involved or affected. In **Switzerland** for example, a [data use policy for the digital education area](#) is being developed with, among other things, ten use cases, which are being accompanied in various contexts. By monitoring the use cases, potentials and challenges are identified, norms of education are discussed, and relevant stakeholders are involved (Educa, 2021).

The challenge of academic honesty in a world driven by AI applications is also a major area of concern at the moment. In the [briefing report 7 Teaching with AI – assessment, feedback and personalisation](#), this area is covered in some detail. It is worth reiterating here for example that some organisations are beginning to put a greater focus on this area both for staff and students. For example, the **International Baccalaureate Organisation** (IBO) puts a strong emphasis on [academic integrity](#) of both teachers and students. Each school has its own academic integrity policy based on the IBO recommendations. As a response to the raising worries of IB educators

worldwide regarding students' usage of AI software, Matt Glanville, Head of Assessment Principles and Practice at IBO, shared in a [blogpost](#) from February 2023 his views on the latest developments of AI. Furthermore, in March 2023, IBO published an official [statement](#) about ChatGPT and AI in assessment and education.

Universities worldwide, concerned by the same academic integrity issue, have started publishing guides on their library websites about how [to reference](#) generative AI and use ChatGPT-like apps in university assignments.

Guidelines on Applying AI Specifically to Education

While the frameworks mentioned above concern AI in general, other guidelines, e.g., [Ethical guidelines](#), [Beijing consensus](#), [Guidance for policy-makers](#) and [Guidance for Generative AI in Education and Research](#) have been published that apply directly to education.

In 2022, the European Commission published the [Ethical guidelines on the use of AI and data in teaching and learning](#) (European Commission, 2022). The purpose is to help educators understand the potential that AI and data usage applications can





have in education and to increase their awareness of the potential risks, so that they can engage positively, critically, and ethically with AI systems and exploit their full potential.

The ethical use of AI and data in teaching, learning, and assessment is based on four key considerations: human agency, fairness, humanity, and justified choice.

- **Human agency** is the capacity of an individual to contribute to society. This is the foundation of autonomy, self-determination, and responsibility.
- **Fairness** refers to all members of a social organisation being treated fairly, including equity, inclusion, non-discrimination, and an equitable distribution of rights and responsibilities.
- Consideration for the people, their identity, integrity, and dignity comprise **humanity**. For a meaningful human connection, we must consider the well-being, safety, social cohesion, meaningful contact, and respect required. It is fundamental to the human-centred approach to AI.
- **Justified choice** refers to the use of knowledge, facts, and data to justify necessary or appropriate collective decisions made by multiple school stakeholders. It necessitates openness and is founded on participatory and collaborative decision-making models, as well as explainability.


Key Requirements for [Trustworthy AI](#) recommendable for any AI system deployed and used in education are human agency and oversight, transparency, diversity, non-discrimination, and fairness, societal and environmental wellbeing and privacy and data governance.

The [Ethical guidelines on the use of AI and data in teaching and learning](#) include guiding questions based on the key requirements for trustworthy AI systems of which the purpose is to facilitate a constructive dialogue on the ethical use of AI in education and training. It could aid the school or educator in formulating pertinent queries and engaging in a productive dialogue with AI system providers or the responsible public bodies.

The International Conference on Artificial Intelligence and Education, held in Beijing in May 2019, led to the adoption of the [Beijing Consensus on Artificial Intelligence and Education](#), the first ever document to offer guidance on how best to harness AI technologies to achieve the Education 2030 Agenda (UNESCO, 2019). Among the recommendations included in this consensus is “ensuring ethical, transparent and auditable use of education data and algorithms”. The parties to the Consensus recommended the following for educators:

“Be cognizant of the dilemmas of balancing between open access to data and data privacy protection. Be mindful of the legal issues and





ethical risks related to data ownership, data privacy and data availability for public good. Be mindful of the importance of adopting principles of ethics-, privacy- and security-by-design.

Test and adopt emerging AI technologies and tools for ensuring teachers' and learners' data privacy protection and data security ... Develop comprehensive data protection laws and regulatory frameworks to guarantee the ethical, non-discriminatory, equitable, transparent and auditable use and reuse of learners' data."

Released in 2021, the [UNESCO guidelines on AI and education for policy makers](#) aim to offer "guidance for policy-makers on how best to leverage the opportunities and address the risks presented by the growing connection between AI and education" (UNESCO, 2021). Some of the policy recommendations are about policies and regulations for equitable, inclusive and ethical use of IA and include to:

- establish and monitor measurable targets to ensure inclusion, diversity and equality in teaching and developing AI services;
- review AI's ability to either alleviate or exaggerate biases;

- create AI applications that are free from gender biases and ensure that the data used for development are gender-sensitive;
- establish data protection laws which make educational data collection and analysis visible, traceable, and auditable by teachers, students and parents;
- investigate options for striking a balance between open access and data privacy;
- facilitate open debates on issues related to AI ethics, data privacy and security, and concerns about AI's negative impact on human rights and gender equality.

Recently, UNESCO also released [Guidance for generative AI in education and research](#) (2023), which aim to support countries in implementing actions to ensure a human-centred vision of these technologies, including by proposing key steps to regulate the use of generative AI in education. Among these steps are notably the adoption and implementation of data protection laws and the definition and enforcement of an age limit for the use of generative AI which should not be below 13 (UNESCO, 2023, p.21)





Data Protection

The right to privacy is a long-established fundamental right at the international level (European Union Agency for Fundamental Rights), enshrined for example in the **International Covenant on Civil and Political Rights and the Convention on the Rights of the Child**. At the European level, it is enshrined in the European Convention on Human Rights (ECHR, Art. 8) and at the level of European Union, in the Charter of Fundamental Rights of the European Union (Art. 7).

Furthermore, other instruments also include provisions on personal data protection, such as the Treaty on the Functioning of the European Union (Art 16) and the EU Charter of Fundamental Rights of the European Union (Art 8), or are even entirely dedicated to the topic, such as the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (Convention 108+), the Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (GDPR), as well as the EU Regulation 2018/1725.

Some non-binding guidelines have also been released, such as the **Council of Europe Guidelines**

[on Children's Data Protection in an Education Setting](#). These guidelines aim to help explain the data protection principles of Convention 108+, previously referred to in this document, and to tackle the challenges in the protection of personal data brought by new technologies and practices, whilst maintaining technologically neutral provisions. The guidelines aim to ensure that the full range of the rights of the child are met as pertains to data protection as a result of interactions with an educational setting, among which are the rights to information, to representation, to participation, and to privacy (Consultative Committee Convention 108, 2020). They should be fully respected and given due consideration for the child's level of maturity and understanding.

In the **United Kingdom, the Open University** UK published a [Policy on Ethical Use of Student Data for Learning Analytics](#), which include 8 principles, among which are defining purpose and boundaries regarding the use of learning analytics, transparency of data collection, and absence of bias. In the **Netherlands**, the SURF foundation published a [guiding paper on how to treat educational data in a privacy conform](#)





[way under the Dutch Data Protection Act](#) in 2017.

This guiding paper on how to use learning analytics (collection and analysis of data, LA) in education under the Dutch Data Protection Act includes a step-by-step plan. It explains what personal data is, the compliance for collection, security disclosure requirements and storage requirements, along with all the obligations institutions must consider before using LA. The providers must be specific about:

- which data is collected;
- why is the data collected;
- how will the data be collected;
- how will the data be used (what will be done with it);
- who had access to the data;
- users right to access data;
- users right to correct or remove data;
- users right to object.

Some interesting points are firstly that the act does not allow for automatic decision-making on a personality profile, as systems using LA can only make recommendations (e.g., the system can grade a student, but not force the student to do

more exercises). All decision-making must be done by human intervention. Furthermore, in the EU processing's agreement cloud services and third parties are required to account for confidentiality, privacy and ownership. The final interesting point is the exception of the law on aggregated data (made anonymous by statistics), which can be used freely.

Some Examples of Cases Regarding Violation of Privacy and Data Protection in Education

Cases have already been brought to court regarding alleged violation of privacy and data protection in education. For example, at the European level, the European Court of Human Rights recognised a violation of Article 8 of the ECHR, in the case of [Antović and Mirković v. Montenegro](#), in 2017. It was initiated by two professors who complained about an invasion of privacy because of video surveillance installed in areas where they taught, on which they stated to have no control over the information collected. The court found that “the camera surveillance had amounted to an interference with the applicants' right to privacy and that the evidence showed that that surveillance had violated the provisions of domestic law. Indeed, the





domestic courts had never even considered any legal justification for the surveillance” (ECHR, 2017).

At the EU level, the European Court of Justice (ECJ), in the case [Nowak](#) in December 2017, stated that Article 2 of the GDPR must be interpreted as meaning that, in circumstances such as in this case, “the written answers submitted by a candidate at a professional examination and any comments made by an examiner with respect to those answers constitute personal data, within the meaning of that provision” (ECJ, 2017).

In France, the administrative court of Marseille cancelled the regional decision to install facial recognition programs in two high schools in 2020. [The decision](#) was partly based on Articles 9 (related to biometric data management), 4. 11 and 7 (related to the notion of consent) of the GDPR. The administrative

court stated that sufficient guarantees should have been put in place to overcome the potential lack of clear and free consent due to the relation of authority between the school and the students. Furthermore, the court also considered that there was not enough proportionality, because it was not demonstrated that usual control such as access badges, and perhaps cameras, were not sufficient enough (TA Marseille, 2020).





Education and Ethics of AI

While AI systems have the potential to bring new opportunities to education and address some of the challenges it is facing, this can only be done by also recognising and addressing the various risks and challenges that can arise with AI. Many rights can be infringed by using AI in education, and the effects of these violations can be very important with long lasting effects, including for (but not limited to) career development and health. For a detailed review on how different human rights can be affected by the use of AI in education, see the report of the Council of Europe on AI and education "[A critical view through the lens of human rights, democracy and the rule of law](#)" (2022).

Some legal protections already exist, notably in terms of privacy and data protection, but it needs to be drastically reinforced, especially regarding ethics of AI, for which no binding instrument exist at the international nor regional level for now. It is encouraging to see that steps are taken at the international and national level to regulate the development and use of AI, however much more needs to be done considering the rapid development of AI and the potential consequences. The recent call for a halt in the development of powerful AI tools is another example of the awareness that exists globally about the urge to address this topic.

The various frameworks and guidelines that currently exist on the topic, while addressing different aspects, often include the following considerations: ethics of AI (including potential bias), AI literacy, gender perspective, societal and environmental wellbeing, the necessity to develop legal frameworks, high sensitivity of data on learners, data protection and the right to privacy, recognition of the roles of teachers, the need to strengthen the research for AI (especially in the public sector), safety, security, transparency, fairness and non-discrimination, proportionality and accountability.

While developing legal frameworks is a prerogative of the states, teaching and learning about AI can already help users to better use and understand AI systems.

We may distinguish between two main perspectives on the use of AI in education. The first one relies on AI-based tools that can be used for different tasks in education, like automatic student assessment, personalised learning, content-creation, etc. The second has to do with teaching and learning about AI, that is AI literacy. But they share a common background in terms of ethics being that adequate training in basic





AI principles provides essential support to students and teachers in judging the response of an AI-based tool from an ethical point of view.

Regarding education with AI-based tools, as explained by Holmes (2023), the ethics behind such systems must address important questions centred on pedagogy, assessments, knowledge, and student and teacher agency. An adequate ethical framework for AI in education needs to be built using learning and human development as a starting point, so it can be used as the core for regulation of AI-based systems used in education. It should be the responsibility of policy makers to create such a regulatory framework, so researchers and commercial actors behind the development of AI-based tools can follow such rules.

AI literacy and digital citizenship are essential topics that should include formal training for responsible use of AI and data-driven technologies, with a critical mindset to be aware of the possible directions and limitations of these systems. Here, an important aim is to help learners to navigate ethical issues related to digital practices, such as that of human autonomy which underpins many of the EU's values. With reliable knowledge about AI from users, learners and educators, the possible unethical response of AI-based tools will be more controlled.

For students at a given age, it will be important to include legal issues such as personal data protection (e.g., GDPR) and privacy, ethical considerations in data collection, storage and use, as well as bias and fairness in AI algorithms. It could also be useful to include examples on AI applications in tools and services, including the use of AI-powered tools for productivity, communication and entertainment, the integration of AI services in custom applications via application programming interfaces (APIs), and the evaluation of AI services in relation to privacy and data security. Similarly, students should be introduced to exploratory data analysis methods using descriptive statistics and data distributions, data visualisation techniques and tools, as well as data-driven decision-making based on automatic analysis.

It must be pointed out that education about AI could follow a developer approach, in the sense that students could act not only as users but also as programmers of simple AI systems. In this realm, it is important that they learn about the ethics and regulations about AI because they will be creating AI-based systems.

To learn more on the topic and on the existing tools helping for teaching about this issue, please refer to [briefing report 3 Use scenarios & practical examples of AI use in education](#).





Recommendations by the Squad

Caution should be a keyword at every level in using AI in education. Students need to be taught their rights and how to protect themselves, teachers need to be cognisant of the range of information collected in the AI tools they use, developers need to guard against undue influence and be aware of potential bias, and finally governmental bodies need to take a firm position with robust legislations to protect their citizens while excising a rigorous approach to their own use of AI in data collection.

In general, we recommend the following learning goals related to AI literacy and ethics:

- Identify and analyse the ethical and environmental opportunities and threats arising from the everyday use of AI.
- Promote a safe, responsible and conscious use of digital tools and technologies related to AI.
- Analyse and understand the human footprint and the influence of risks in automated decision-making processes.
- Identify and evaluate the ethical and policy implications of the design and use of AI systems, including fairness, bias, discrimination and accountability.
- Critically analyse the potential of AI to improve peoples' quality of life, assessing its operability in different social, economic and cultural contexts.
- Know and understand the risks and benefits of AI in different areas, such as health, security and privacy.



Members of the EDEH squad on artificial intelligence in education who dedicated time for this briefing report: Elise Rondin, Francisco Bellas, Martina Weber, Petra Bevek, Bertine van Deyzen, Jessica Niewint-Gori, Cristina Obae, Anne Gilleran and Lidija Kralj.

