An initiative of the European Commission

Influence of AI on Governance in Education

Briefing report No. 5

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

EUROPEAN DIGITAL EDUCATION HUB

The European Digital Education Hub (EDEH) is an initiative of the European Commission, funded by the Erasmus+ programme (2021-2027) and operated by a consortium of 11 organisations under a service contract with the European Education and Culture Executive Agency (EACEA).

Content

Introduction

Education governance refers to **how decision making happens in education systems** and how education systems allocate roles and responsibilities, determine priorities and designs, and carry out education policies and programmes (OECD, 2019). From an education governance point of view, it is increasingly important to explore and discuss the possibilities, risks and limits of artificial intelligence (AI) in education. Observing the institutionalisation of new education governance practices that emerge as a result of the integration of digital technologies into education is necessary in order to share best practices and gain knowledge. To discuss these new governance practices, the UNESCO guidance for policy makers for AI in education and European Parliament proposal of AI Act (adopted text, June 2023) are taken into consideration.

A number of AI tools for educational purposes are already in use (see Briefing report No.3 "Use Scenarios and practical examples of AI use in Education"). Many positive examples for effective use start to emerge, however, there are also many concerns for responsible adoption, such as the lack of strategies to specify measures that are conducive to effective use of AI for educational purposes. There is a need for establishing an integrated education governance package for AI that encompasses educational reform, ensuring inclusive, equitable and ethical use of AI. Policies and strategies for using AI in education are central to maximising AI's benefits and mitigating its potential risks as a new tool to accelerate the progress towards the achievement of the UN's sustainable development goal 4 (SDG 4) – **Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.**

Risk-Based Approach

In June 2023, the European Commission has proposed, and the European Parliament has adopted the text for the legal framework on AI which addresses the risks of AI ensuring that Europeans can trust what AI has to offer. While most AI systems pose limited to no risk and can contribute to solving many societal challenges, certain AI systems create risks that need to be addressed to avoid undesirable outcomes.

"Depending on the circumstances regarding its specific application and use, as well as the level of technological development, artificial intelligence may generate risks and cause harm to public or private interests and fundamental rights of natural persons that are protected by Union law. Such harm might be material or immaterial, including physical, psychological, societal or economic harm" (AI Act proposal, EP, June 2023). The proposed text also points out some potential risk of AI in education which will be highlighted below.

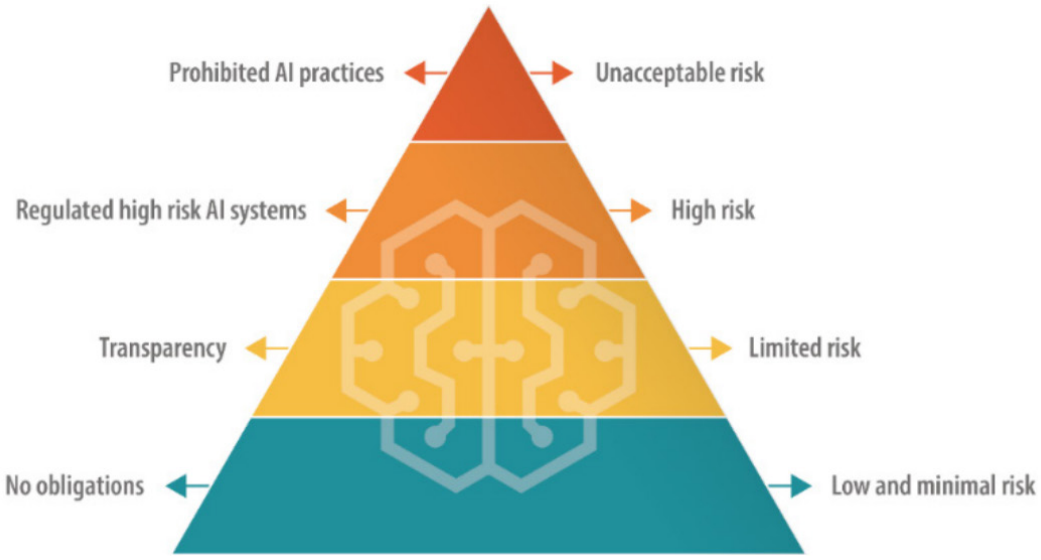

Figure 1 Illustration from European Parliament briefing on AI Act (January 2022)

Unacceptable Risk

All AI systems considered a clear threat to the safety, livelihoods and rights of people will be banned. Prohibited are among others: crime prediction based on the profiling of individuals, 'real-time' remote biometric identification systems in publicly accessible spaces and 'post' remote biometric identification systems, creation of facial recognition technology databases through the scraping of facial images from the internet or CCTV footage and **inferring the emotions of individuals** in certain areas, such as in workplace, by **education institutions**, law enforcement or border management (AI Act proposal, EP, June 2023).

High Risk

AI systems identified as high-risk include AI technology used in critical infrastructures, 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).

High-risk AI systems should have adequate risk assessment and mitigation systems, high-quality datasets, logging of activity, detailed documentation, clear and adequate information to the user, appropriate human oversight measures, and high level of robustness, security and accuracy (EC, 2022). High-risk AI systems shall be tested for the purposes of identifying the most appropriate and targeted risk management measures and weighing any such measures against the potential benefits and intended goals of the system. Testing shall ensure that high-risk AI systems perform consistently for their intended purpose and they are in compliance with the requirements set out in the AI Act proposal (EP, June 2023).

"**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, EP, June 2023).

"Hight-risk AI systems in education and vocational training are:
(a) AI systems intended to be used for the purpose of determining access or materially influence decisions on admission or assigning natural persons to educational and vocational training institutions;
(b) AI systems intended to be used for the purpose of assessing students in educational and vocational training institutions and for assessing participants in tests commonly required for admission to those institutions;

- **(b a) systems** intended to be used for the purpose of assessing the appropriate level of education for an individual and materially influencing the level of education and vocational training that individual will receive or will be able to access.
- **(b b) AI systems** intended to be used for monitoring and detecting prohibited behaviour of students during tests in the context of/within education and vocational training institutions." (AI Act proposal, EP, June 2023).

Limited Risk

Limited risk refers to AI systems with specific transparency obligations. For example, when using AI systems such as chatbots, users should be aware that they are interacting with a machine so they can take an

informed decision to continue or step back (EC, 2022).

Minimal or no Risk

The AI Act proposal allows the free use of minimal-risk AI. This includes applications such as AI-enabled video games or spam filters.

How does it all work in practice for providers of high risk AI systems?

Figure 2 Illustration from Regulatory framework proposal on artificial intelligence (EC, 2022)

AI applications should remain trustworthy even after they have been placed on the market. This requires ongoing quality and risk management by providers.

Emerging Practices and Benefit-Risk Assessment for AI in Education

The application of AI in education has developed in multiple directions as is already mentioned in Briefing report No.1 "Teacher's competencies", and No. 3 "Use scenarios & practical examples of AI use in education". Importantly, the introduction of AI into education also put and accent on issues of pedagogy, organisational structures, access, ethics, equity, and sustainability.

Furthermore, if the potential of AI to support education for sustainable development is to be fully realised, all the possible benefits of the tools need to be identified and leveraged, and the risks acknowledged and mitigated. Consequently, the ways in which education is organised also need to be continuously reviewed, which might suggest a fundamental reshaping of education's core foundations.

To help and support educational systems to respond to these complex challenges, UNESCO in its Beijing Consensus on Artificial Intelligence and Education determines more than thirty policy recommendations related to the questions:

- How can AI be leveraged to enhance education?
- How can the ethical, inclusive, and equitable use of AI in education be ensured?
- How can education prepare humans to live and work with AI?

From the perspective of governance of AI systems in education, and through a prism of EU ethical principles, some examples of guiding questions are proposed in a JRC publication on Emerging technologies and the teaching profession: Ethical and pedagogical considerations based on near-future scenarios (2020):

- What procedures and policies are in place to support humans in the working environment, and the aim of creating meaningful work?
- What procedures and policies are in place to ensure that the AI systems cannot cause or exacerbate adverse impacts due to asymmetries of power or information, such as between employers and employees, businesses and consumers or governments and citizens?

- What protocols are in place to respond to and to prevent harm? What early warning systems are there that can trigger action if harm may be occurring?
- What procedures and policies are there to ensure that AI systems positively address rather than exacerbate inequity, discrimination and prejudice in education? What evidence is there that an AI system can be used to address equity concerns in schools?
- How will those in governance or procurement positions ensure genuine traceability, verifiability, non-deception and honesty, and intelligibility of AI systems prior to purchase and during implementation?
- How will transparency be operationalised if harm occurs?
- Do policy-makers, procurement officers, and school leaders have access to appropriate independent technical expertise to explain and advise on AI systems?
- Is there a rigorous process for seeking parental consent and student assent before systems are deployed?

More Guiding questions, from a teachers' and school leaders' perspective are available in the Ethical guidelines on the use of artificial intelligence and data in teaching and learning for educators (2022, p. 19-25) to facilitate the evaluation of an AI system prior to its installation in a school or during its use. The inquiries may be directed to the educators themselves, the decision-makers at the management level, or the system providers. The queries can also be used to guide discussions with students, parents, and other members of the school community.

AI Strategies and Planning the Use of AI in Education

According to the European Commission's AI Watch, almost all European countries have an AI strategy, 23 out of 27 member states as stated in the 2022 edition of AI Watch.

The European Schoolnet report Artificial Intelligence. Role in K12 Education (2021) mentions that national policies addressing AI and education developments are diverse. Education is mostly mentioned in general AI strategies or AI is mentioned in education or digital strategies, while a thematic approach focusing on AI and education is rare.

According to the 2022 edition of AI Watch, AI strategies usually dedicate a section related to AI skills. The focus is mainly put onto measures for higher education institutions and the reskilling of the working age population, only briefly mentioning primary and secondary education. Few countries envisage policy measures for pre-school education and set out mechanisms for the impact assessment of skill policies. Countries are working on several models for reskilling and upskilling and the introduction of AI skills in formal and informal education, but specific measures targeting assessment of AI use in education are not present.

The AI Act proposal (June 2023) also mentions the importance of AI literacy mandating that providers and deployers of AI systems ensure a sufficient level of AI literacy among their staff and other individuals dealing with AI systems on their behalf, taking into account their technical knowledge, experience, education, and training, as well as the context in which the AI systems will be used. The proposal explains that 'AI literacy' refers to the skills, knowledge, and understanding that enable various stakeholders to deploy AI systems in an informed manner and to know about the opportunities, risks, and potential harms of AI. The AI Act proposal emphasises the importance of promoting the development of a sufficient level of AI literacy across all sectors of society, for people of all ages, including women and girls.

In the Beijing Consensus on Artificial Intelligence and Education, UNESCO suggests four strategic targets which should be covered in AI policies:

- ensuring the inclusive and equitable use of AI in education;
- leveraging AI to enhance education and learning;
- promoting the development of skills for life in the age of AI, including teaching how AI works and its implications for humanity;
- safeguarding the transparent and auditable use of education data.

The planning should be interdisciplinary and governance inter-sectoral. The policies need to be developed on equitable, inclusive, and ethical use of AI planning to use AI for education management, teaching, learning, and assessment. Pilot testing, monitoring, and evaluation, and building an evidence base as well as fostering local AI innovations for education should be carried out.

The education policies should be tailored to equip current and future generations with the necessary tools to harness AI for sustainable development and ensure that it operates based on **human rights**, as a shared set of values and ethics for the benefit of all humanity. When adopting the education policies for AI, the steering wheel must be a humanistic approach. AI education policy development and practices should focus on protecting human rights and equipping people with the values and skills needed for sustainable development and effective human-machine collaboration in life, learning and work (UNESCO).

AI implementation in education should be based on **system-wide readiness** and cost-value assessment including assessment of infrastructure, internet connectivity, the availability of data, AI tools, local AI talent, the skills of key policy implementers, and stakeholders' awareness. The Government AI Readiness Index 2022 presents readiness of governments to implement AI in the delivery of public services, monitoring 39 indicators across ten dimensions, organised in three pillars: the government pillar; the technology sector pillar, and the data and infrastructure pillar.

Interdisciplinary and multistakeholder expertise should be mobilised to inform policy planning and build the capacities of policy-makers. Inter-sectoral governance and coordination mechanisms, open and iterative cycle composed of key steps in planning, implementing, monitoring and updating policy need to be set up (UNESCO).

Pilot testing, monitoring and evaluation, and building an evidence-based approach of applying AI in education are necessary steps, while also putting efforts in strengthening **research and evaluation** in the field of AI and education and foster local AI innovations for education. In May 2023, the International Research Centre On Artificial Intelligence published the IRCAI Global Top 100 Report 2022 for which they evaluated projects across four key criteria: scientific maturity and use of AI tools, impact on relevant SDGs, market readiness, and ethical ramifications, with particular attention to ethical and rights-based awareness of the applicants in their proposed AI solutions.

The Ethical guidelines on the use of artificial intelligence and data in teaching and learning for educators (2022, p. 26- 27) suggest using an incremental approach to **gradually introduce AI tools into school contexts** and to constantly monitor the societal effects that can emerge, leaving open the possibility to step back when unintended consequences occur. Several steps are proposed for planning effective use of AI and data in school:

1. reviewing current AI systems and data use;
2. initiating policies and procedures;
3. carrying out a pilot of the AI system;
4. collaborating with the AI system providers;
5. monitoring the operation of the AI system and evaluate the risks;
6. raising awareness and community engagement (discussing and collaborating with colleagues and other schools, communicating with parents, learners and the school community and keeping them up to date).

Recommendations by the Squad

With AI and the associated data, new norms and new governance models emerge, and new actors enter the education sector while others lose their value in the system. Although national authorities are reacting quickly establishing or improving their AI strategies, it is difficult at the moment to have a clear picture of what this virtual AI-based ecosystem will look like, what governance it will have, and what actors will be involved, but four common areas of concern emerge from the national and regional policies:

- the importance of governance for data and privacy;
- the importance of openness to ensure equal universal access and promote transparency;
- curriculum innovation that can address the potential and implications of AI;
- financial support for the effective implementation of AI.

The primary purpose of applying AI in education should be to enhance learning, enabling every learner to develop their individual potential, and policies should reflect and support it. A comprehensive AI strategy is recommended covering interdisciplinarity, humanity, ethics, scalability and sustainability, responsibility, equity and lifelong learning for all.

Members of the EDEH squad on artificial intelligence in education who dedicated time for this briefing report: Gordana Janakievska, Riina Vuorikari, Yann-Aël Le Borgne, Martina Weber, Cristina Obae, Jessica Niewint-Gori, Anne Gilleran and Lidija Kralj.