

Regulating artificial intelligence in the education domain: a general approach

Ilkka Tuomi

While many proposals exist on how to regulate and govern artificial intelligence (AI), very few address education and human development. I will argue that the education domain poses special challenges for AI regulation and that all regulatory frameworks for AI should be grounded on fundamental concerns related to education and learning.

To understand how and why this is necessary, it is important to know what the current regulatory initiatives are and how they differ from what is needed in the education domain. In this policy-oriented brief, I suggest some entry points from which AI regulation in the education domain could be developed.

Current approaches in AI regulation

Although there have been many proposals on how to govern and regulate AI, two prominent ones are the European Union (EU) AI Act and the US Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. Both require that the risks of AI systems are managed and that the developers of high-risk systems must provide adequate information on their systems for this. Although the EU AI Act builds on an ethical framework with human rights at its foundation, its key objective is to create unified regulation that allows the development and use of AI systems without unnecessary restrictions within the single EU market. In the United States of America, a single market already exists. Partly because of that, the US approach emphasises risks to national security, critical infrastructure and national competitiveness. Both aim to balance risks with opportunities for innovation and technical progress.

The US Executive Order has launched many studies and reports that aim at specifying what needs to be done to govern AI.¹ It addresses education mainly indirectly, as an instrument to create technical competences in AI, both by attracting and developing AI talent and as a tool to help the workforce adapt to the labour market changes brought by AI. Section 8 of the Executive Order encourages independent regulatory agencies, as they deem appropriate, to consider using their full range of authorities to protect American consumers, patients, passengers and students. In this context, the secretary of education is ordered to develop resources, policies and guidance regarding AI that shall address safe, responsible and non-discriminatory uses of AI in education. These resources shall also include the development of an “AI toolkit” for education leaders. The toolkit is to implement recommendations from the Department of Education’s “Artificial Intelligence and the Future of Teaching and Learning” report, published in 2023, which suggests an appropriate human review of AI decisions, designing for trust and safety, and alignment with privacy-related laws and regulations in the educational context.²

1. US Presidential Action (30 October 2023), Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, available at www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/, accessed 20 June 2024.
2. US Department of Education, Office of Educational Technology (2023), “Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations”.

Whereas the US approach is strongly focused on threats to national security and competitiveness, the EU AI Act aims to build trust that facilitates wide adoption of AI. The AI Act proposal aimed “to give people and other users the confidence to embrace AI-based solutions, while encouraging businesses to develop them”.³ Instead of national interests, the proposal from the European Commission focused on the concerns of citizens that need to be addressed to create a level playing field for AI developers and providers. Education was defined as one of the potential areas where AI systems could create high risks for citizens, specifically if AI systems can influence access to education and future life opportunities. As part of its Digital Education Action Plan, the European Commission also developed guidelines for the ethical use of AI and data in education aimed at teachers.⁴ The guidelines provide general suggestions on how to approach the use of AI in the educational sector and could be considered as soft regulation through capacity development.

In general, the proposals for regulating AI usually suggest transparency and disclosure, registration, licensing and audit mechanisms. The requirement for disclosure includes documentation that adequately informs regulators and citizens about potential risks and their management. This could include documentation on the appropriate uses and known risks, as well as documenting quality control and risk management systems. Transparency also includes the principle that humans should be informed when they are interacting with AI systems and, in the context of generative AI, specifically the requirement that AI-generated content is labelled as such. At a more technical level, the use of copyrighted content for the training of AI systems needs to be documented and various watermarking systems have been proposed for generative AI. Registration of high-risk AI systems, in turn, allows regulators to know and control what kinds of AI systems are put in use, and this can be enforced, for example, by certificates or licences. Granting such licences can further require evaluation and audits of the promised functionality by certification authorities or third-party auditors, including the assessment of alignment with existing regulations and approaches used to mitigate known harms.

As these examples highlight, existing regulatory approaches aim at secure, safe and trustworthy AI systems. In line with this, the recent United Nations General Assembly Recommendation on AI regulation and governance suggests several actions that the member states could take to realise the potential of AI in supporting the Sustainable Development Goals and to limit the risks of AI. The recommendation encourages the member states to consider structural impediments and digital divides across and within countries, as well as the potential of AI to promote digital transformation, peace, and the enjoyment of human rights and fundamental freedoms when regulatory frameworks are being developed.⁵ The recommendation also notes the importance of research on the impacts on the labour market, and tools for assessing the reliability and provenance of AI-generated content.

In the EU AI Act negotiations, the European Parliament strongly emphasised the importance of protecting democracy, the rule of law and the environment. The final compromise text notes that – except for an introductory high-level statement – the regulation text only addresses risks to health, safety and fundamental rights. The fundamental rights impact assessment highlighted by the European Parliament will also be substantially limited. It is expected to be based on a standard questionnaire provided by the European Commission AI Office and only covers deployers that are bodies governed by public law, private actors providing public services, and some banking and insurance service providers.⁶

These developments are important and they will help to promote human rights, democracy and the rule of law. It is, however, useful to ask whether such regulations and encouragements for action address the needs in the educational sector in a satisfactory manner.

3. European Commission (2021), Proposal for a Regulation of the European Parliament and of the Council: Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021) 206 final), Section 1.1, available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>, accessed 20 June 2024.

4. European Commission (2022), *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*, available at <https://op.europa.eu/en/publication-detail/-/publication/d81a0d54-5348-11ed-92ed-01aa75ed71a1/language-en>, accessed 20 June 2024.

5. United Nations (2024), Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development (UN General Assembly Resolution A/78/L.49), available at <https://undocs.org/A/78/L.49>, accessed 20 June 2024.

6. Council of the European Union, (2024), Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts – Analysis of the final compromise text with a view to agreement (5662/24).

Education is different

Education is of fundamental importance in a world where AI-based systems are widely used. Education, however, has received relatively minor attention in regulatory debates, beyond the frequent declarations that citizens, educators, workers, policy makers and computer programmers need more AI-related competences to survive and prosper in the emerging world.

Partly, this lack of attention results from the fact that, for historical reasons, education is governed by local authorities. Because of this, federal and EU-level education policies have predominantly focused on those aspects of education that are directly related to general policy objectives, such as promoting economic growth, employment and innovation. This has led to a very instrumental view on the role of education in society.

A broader view on the role of education is needed to understand why and how AI should be regulated in the educational sector.

Education researchers⁷ commonly distinguish three main functions or purposes of education in society. The first one is qualification. Here, education is concerned with the transmission and acquisition of knowledge, skills, dispositions and understandings that qualify people to do certain things. The second function is socialisation. Through education, people become part of existing traditions, cultures, ways of doing and ways of being. This is sometimes called the “hidden curriculum”. The third function of education in society can be called subjectification. This has to do with how education contributes to how we can exist as human subjects. In the influential Delors report on lifelong learning,⁸ these purposes were defined as “learning to know”, “learning to do”, “learning to be” and “learning to be together”. The Recommendation of the Committee of Ministers on the public responsibility for higher education and research⁹ adds the development and maintenance of society’s knowledge base as one important purpose of education and related research.

The use of AI in education will have important implications for all these purposes of education. In general, extending beyond the industrial-age interest in creating useful skills and knowledge for employment, the broader purposes of education highlight the role of education in human development. In the context of emerging technologies in education, it has therefore been suggested that the sustainable objective for education will be the development of human agency.¹⁰ Agency refers here to an individual’s capability for action and characterises what a person is free to do and achieve in pursuit of goals and values the person regards as important.¹¹ Agency requires skills, knowledge, experience and meta-cognitive competence components, but also technical infrastructures and social and cultural resources, including norms and laws.

AI, together with the other key technologies of the “next internet”, will generate a new infrastructure for knowing and action, and this will have implications for human development.¹² For several key theorists of learning, learning is a process that changes the ways in which we think and perceive the world. This, in turn, changes what we do and what we are able to do.¹³ One could, therefore, define education as a process that makes us able to participate in – and change – social and cultural life.

This view also underpins the Common European Framework of Reference for Languages (CEFR)¹⁴ that sees learners as social agents who actively exert their agency in shaping and using social resources, including language.¹⁵

7. For example, Biesta G. J. J. (2010), *Good Education in an Age of Measurement: Ethics, Politics, Democracy*, Routledge.

8. Delors J. (1996), *Learning: the treasure within*, UNESCO, available at <https://unesdoc.unesco.org/ark:/48223/pf0000109590>, accessed 20 June 2024.

9. Recommendation CM/Rec(2007)6 of the Committee of Ministers to member states on the public responsibility for higher education and research, available at <https://rm.coe.int/16805d5dae>, accessed 20 June 2024.

10. Tuomi I., Cachia R. and Villar-Onrubia D. (2023), “On the futures of technology in education: Emerging trends and policy challenges”, JRC Science for Policy Report, European Commission, available at <https://doi.org/10.2760/079734>, accessed 20 June 2024.

11. Sen A. (1993), “Capability and well-being”, in Nussbaum M. C. and Sen A. (eds), *The Quality of Life*, pp. 30-53, Clarendon Press.

12. Tuomi I., Cachia R. and Villar-Onrubia D. (2023), “On the futures of technology in education: Emerging trends and policy challenges”, JRC Science for Policy Report, European Commission, available at <https://doi.org/10.2760/079734>, accessed 20 June 2024.

13. Beyond philosopher and educational reformer John Dewey, this link between activity and learning has been a specific focus in the Vygotskian cultural-historical activity theory. For psychologist Lev Vygotsky, “higher levels” of thought emerge when children learn to use culturally accumulated conceptual systems in their thinking. These, in turn, reflect historically developed forms of social practice and division of labour. These cultural systems of thinking, according to Vygotsky, cannot emerge from simple maturation but require education that guides a developing child towards adult forms of thinking. This is in contrast to psychologist Jean Piaget, who also emphasised the link between cognition and action, but mainly viewed cognitive development as maturation.

14. Council of Europe (2020), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, Council of Europe Publishing, available at <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>, accessed 20 June 2024.

15. Piccardo E. and North B. (2020), “CEFR Key Concepts: The user/learner as a social agent”, Council of Europe, available at <https://rm.coe.int/the-user-learner-as-a-social-agent/1680a862ac>, accessed 20 June 2024.

Such a view on education extends beyond a simple utilitarian view. Specifically, education cannot only be understood as something that produces given and predefined consequences.¹⁶ This also means that the implicit consequentialist ethics¹⁷ that underpin present debates on AI regulation do not provide an adequate foundation for understanding what needs to be regulated when AI is used in education. In a more developmental view, the regulation of AI cannot only be about making AI systems safe or about protecting established universal principles. More fundamentally, AI regulation needs to be considered as a tool that is used to ensure that AI systems support and do not hinder social and individual development.

Instead of harm, focus on ethics and development

Consequentialist ethics is a challenging starting point in a world where innovation creates new realities. Many influential contributions to AI regulation and ethics state that AI should be beneficial and that it should “do no harm”. Such consequentialist approaches often make unrealistic epistemic assumptions. It has been known since antiquity that causal chains are infinite. More recently, research on innovation, technology development and knowledge creation has pointed out that many of the beneficial and harmful consequences of innovation and learning are unintended and unpredictable. This suggests that ethical traditions that put more emphasis on the openness of the future and responsible action can be more suitable in the educational domain. When we do not have enough knowledge to avoid harm, we need to make responsible and justified choices and learn from mistakes.

Such a view underpins a socio-developmental approach to ethics of AI in education. In this framework, the developmental aspect becomes reflected in the impact of AI systems and their use in the development of human agency. Individual and social differences, in turn, become reflected in social fairness. The need to negotiate and agree on collective action forms the third pillar of ethics, that of justified social choice.¹⁸

To justify the use of an AI system for teaching, for example, means that there is an accepted process for making trade-offs that results from the reallocation of resources, including money and teacher and student time. To justify such a change, some justification would be needed, and this could, for example, consist of acceptable evidence that learning, education and individual development could benefit from the change. What types of knowledge and evidence, and whose voices are heard, therefore, become important questions for regulation.

Whereas many AI ethics recommendations have tried to formulate universal principles that would guide the development and use of AI systems, a more socio-developmental approach suggests a process- and learning-oriented approach. Such an approach also requires new thinking. Technology regulation frequently focuses on the characteristics of technology, as if technical artifacts would need to be regulated instead of their use and impact. AI systems are often general-purpose technologies and regulation must be based on their use. Their use is unpredictable as new forms of use are continuously invented and the impact always depends on the social context where technology is deployed.

Assuming that we cannot predict and control the future, the question about regulating AI shows up in a new light. Effective regulation becomes a question of shaping the paths of development for the emerging uses of technology. Prediction is replaced by policy.

Instead of providing guardrails or safeguarding people and learners from expected harm, the aim of regulation should be on pushing technological development in directions that support education, in all its main social functions.

One way of doing this is to make real progress attractive for technology developers. This, of course, requires having an idea of what progress in education means. A shared view on progress in education requires a future-oriented view. Regulation, therefore, must express our aspirations and objectives for the future of education and avoid formalising the past.

Assess impact on agency and its development

A responsible choice of adopting AI tools in education is based on answering the question: does the use of AI increase the development of human agency and does it do this with social fairness? In specific cases and

16. Tuomi I. (2023), “Beyond mastery: Toward a broader understanding of AI in education”, *International Journal of Artificial Intelligence in Education*, Vol. 34, pp. 20-30, available at <https://doi.org/10.1007/s40593-023-00343-4>, accessed 20 June 2024.

17. Consequentialism holds that the anticipated outcomes of an action are the basis for judging their moral or ethical value.

18. Tuomi I. (2023), “A framework for socio-developmental ethics in educational AI”, *Proceedings of the 56th Hawaii International Conference on System Sciences*, 6208-6217, available at <https://hdl.handle.net/10125/103386>, accessed 20 June 2024.

contexts, there can be arguments for and against, and these need to be negotiated. The negotiation process itself is a process of collective knowledge creation and social learning. It, therefore, cannot be replaced by a predefined checklist. There is no algorithm that can make optimal choices for us in an open world.

From a regulatory point of view, this suggests that beyond a mandatory fundamental rights impact assessment, a more developmental agency development impact assessment would be useful. Such an assessment, of course, would be relevant also beyond formal education as lifelong learning and informal learning can be expected to be increasingly important when AI technologies will be widely used. In fact, one core distinguishing characteristic of AI is its capability to be agentic in ways that were not possible for traditional technologies. Agency, itself, is becoming distributed in new ways across humans and AI systems. Education researchers have explored AI-augmented learning and thinking since the 1980s;¹⁹ at present, we still have very limited understanding about the cognitive, social and cultural impacts of AI technologies.

In such a developmental context, we can ask what regulation and governance approaches are needed to make AI systems better. Instead of focusing on regulation that restricts technology development and use, we can ask what is needed to promote development, understood as social and human development. Ethics of education and human development are focused on this question and therefore could provide the foundation for regulating AI also more generally.

Regulate through policy learning and capability expansion

A process-oriented view on regulation can set rules for how social choices are made and what types of evidence should be considered. The requirement to assess fundamental rights impact defines a set of considerations that need to be examined. The impossibility of knowing the impacts, however, suggests that policy learning becomes important for AI regulation.

In the context of educational institutions, this begs the question about procedures and responsibilities. An important element in policy learning is feedback from observed problems. Although there are now several proposals for registries that record AI-related incidents, education and human development-specific incidences and concerns are not well covered. A registry of observed challenges and their solutions would be an important part of the policy learning process. Furthermore, many important concerns in the educational domain cannot be understood as “incidents” as they are related to long-term societal impact.²⁰

Organisational learning, however, requires absorptive capacity.²¹ Without adequate competences, educational organisations will not be able to internalise and appropriate new knowledge. This suggests that competence development is a key element in regulating AI. For example, the United Nations Educational, Scientific and Cultural Organization (UNESCO) AI Competence Framework for Teachers, currently being developed, defines several competence elements that teachers need to develop to be able to appropriately use AI. Hard regulation only works if it is complemented with soft regulation based on competence development.

Manage power shifts

The complexity of AI and advanced information systems amplify the voices of technology experts. Whereas education and education policies have traditionally been shaped by local policy makers and stakeholders, these policies are now increasingly being formed by global technology providers. At a very fundamental level, AI providers are actively redefining what intelligence, learning and education mean in policy discourses.²²

An important element in this power shift is access to data on learning. Fine-grained data on learning are needed to develop AI for teaching and learning, and for advancing research on learning. At present, such data on learning are available only to the large platforms and their customers who can pay for it, and only very few large education-specific technology providers have access to data that can effectively be used to develop data-driven AI systems. For example, according to Carnegie Learning, its recently announced LiveHint AI math tutor,

19. For example, Salomon G. (ed.), (1993). *Distributed cognitions: Psychological and educational considerations*, Cambridge University Press.

20. Tuomi I. (2022), “Artificial intelligence, 21st century competences, and socio-emotional learning in education: More than high-risk?”, *European Journal of Education*, 57(4), pp. 601-619, available at <https://doi.org/10.1111/ejed.12531>, accessed 20 June 2024.

21. Cohen W. M. and Levinthal D. A. (1990), “Absorptive capacity: A new perspective on learning and innovation”, *Administrative Science Quarterly*, 35, pp. 128-152.

22. Blikstein P., Zheng Y. and Zhou K. Z. (2022). “Ceci n’est pas une école: Discourses of artificial intelligence in education through the lens of semiotic analytics”, *European Journal of Education*, 57(4), pp. 571-583.

which combines knowledge-based AI with a data-driven large language model, is built by drawing insights from the experiences of 5.5 million students tackling over 1.2 billion math problems.²³

Data-driven AI is based on data and access to data on learning will greatly influence how AI systems will be developed in the future. An important development in this area is the EU Data Act that potentially moves power over data back to the users.²⁴

Hard, soft or smart regulation?

As the EU AI Act aims to protect fundamental rights, democracy and the rule of law, one must ask whether there remains a need for further regulation in the educational domain. In theory, education has been defined as a high-risk area and the European Commission could add new high-risk uses to address emerging risks as they become better understood.

From the discussion above, it should be clear that AI regulation in the educational domain must be rooted in concerns that are specific to education. Education is about human, social and cultural development. AI is changing the infrastructures of knowing and learning, and the emerging opportunities and risks can only be understood by revisiting what we know about human development. Education makes us able to be competent participants in society and culture, and agents of their change. AI will expand human agency, which in the process also becomes distributed in new ways across human and technological actors. Much remains to be done to guide education and AI towards directions that we have reason to value.

As technical developments are now driven by large businesses that are fundamentally motivated by profit, law can provide incentives and sanctions that align commercial interests with future-oriented education policy. Without a shared view on the role of education, it is, however, impossible to have a shared normative order that becomes expressed in laws and regulations. Such a normative order must be aligned with ethics of education that goes beyond purely instrumental views. Education is more than about manufacturing skilled workers for future employment.

For educators and policy makers, law can become an important co-ordinating tool and signpost. Creating regulation for the educational uses of AI, therefore, is an important opportunity to express the values and norms that underpin education. Educators are the key stakeholders in this process. Pedagogic and andragogic knowledge needs to be at the centre, and economic concerns that often dominate the industrial-age debates on the future of education should become secondary.

AI regulation in education also requires substantial and procedural legitimacy. Substantially, it must be aligned with the ethics of the key stakeholders. Procedurally, laws and regulations must be shaped in a process that the stakeholders consider legitimate. As technology is rapidly changing, existing norms require reassessment. It is not yet clear what concepts can provide an adequate normative foundation. Also because of this, stakeholders need to be involved in a joint creation of shared frameworks that can underpin future legislation and AI governance models.

23. Business Wire (30 November 2023), "Carnegie Learning announces LiveHint AI™", available at www.businesswire.com/news/home/20231130974040/en/Carnegie-Learning-Announces-LiveHint-AI%E2%84%A2, accessed 20 June 2024.

24. European Council (27 November 2023), Data Act: Council adopts new law on fair access to and use of data, available at www.consilium.europa.eu/en/press/press-releases/2023/11/27/data-act-council-adopts-new-law-on-fair-access-to-and-use-of-data/, accessed 20 June 2024.