on_education
Journal for Research and Debate

taming algorithms

Tobias Röhl

The introduction of artificial intelligence (AI) and other tools, based on algorithmic decision-making in education, not only provides opportunities but can also lead to ethical problems, such as algorithmic bias and a deskilling of teachers. In this essay I will show how these risks can be mitigated.

Keywords: AIED, algorithms, artificial intelligence, education, ethics

Students were furious. All over the UK they rallied against a common foe. Their foe reduced their grades from As to Bs, from Bs to Cs and even worse. Many felt that they were treated unjustly. Holding up placards with statements such as “teachers know my potential, algorithms do not” their discontent was aimed at an algorithm employed by the Department of Education in an attempt to battle “grade inflation”.*1* This essay explores the unwanted effects of algorithms and considers how they can be tamed.

In order to understand the relevance of these issues, one needs to position algorithms in the context of current debates and developments. Machine learning allows the automation of algorithmic-decision making and the constant refinement of data-based tools. Consequently, there is a growing interest in artificial intelligence (AI) and other tools, based on machine learning. They are at the core of current imaginaries of the future and are painted as a disruptive force, bringing about considerable societal transformations. Artificial intelligence in education (AIED) accordingly promises to ”unleash intelligence“ (Luckin et al., 2017) and to result in “pushing the frontiers“ (OECD, 2021). AIED can assist teachers in several domains: “Intelligent tutoring systems” (ITS) automatically select tasks and provide individual feedback, based on the performance of students (Clement et al., 2015), connecting each student with their “superteacher” (TA-SWISS, 2020, p. 14); “learning analytics“ (e.g., by Microsoft) give a data-driven overview of the strengths and weaknesses of learners, classes, and schools (Kop et al., 2017) and “robo-grading“ promises the objective and efficient evaluation of tests and assignments (Foltz, 2015). The hope is that teachers turn from “sage on the stage” to “guide on the side” (Susskind & Susskind, 2015, p. 60) providing their informed and individual assistance to students only when required.

At the same time, some authors warn of AI’s shortcomings and highlight the ethical problems that could arise. “Algorithmic bias” (Akter et al., 2021) is now a widely recognized problem of algorithmic decision-making. The datasets used to train AI, based on machine learning, contain the social bias and prejudices inherent to human practice. A popular example is Amazon’s failed attempt to introduce AI as a recruiting tool.*2* Soon after its introduction, it became clear that the algorithm was biased against applications from women, since it was based on the data of previous applications according to which women were supposedly less likely to be interviewed and recruited

than men.

The aforementioned algorithm which infuriated students is another instance of algorithmic bias. Due to the pandemic, the UK government decided to cancel the A-level exams in 2020. Instead, teachers were asked to estimate the results of their students, which consequently lead to average results being much higher than in previous years. An algorithm was employed to take care of the situation and correct the results based on data from previous years. Unfortunately, the algorithm operated partly on the basis of average results from a student's school. Consequently, the grades of students from private schools were reduced to a lesser extent than those of students from state schools. Students from state schools felt that they were treated unfairly. Indeed, the algorithm denied the existence of outliers, such as exceptional students in schools who, on average, had performed poorly in the past. Similarly, the results in schools that had undergone reform could have been better in the year that the grades were estimated. However, soon after the students went to the streets, the decision to use the algorithm was retracted.

Algorithmic bias is, however, not the only problem casting doubt on the notion of objective and flawless algorithms employed in education. On a more fundamental level, AIED, such as intelligent tutoring systems can also be viewed as "achievement technology" (Chang, 2019, p. 40) promoting the concept of learning as an individual performance (Macgilchrist et al., 2020). Each student learns on his/her own, at his/her own pace, with individual tasks selected for him/her. Such individual concepts of learning favor students that are already highly motivated and self-organized (Selwyn, 2019, p. 95). At the same time, notions of collaboration and participation in a collective are neglected. However, not only students are affected by AIED. There is a fear of a deskilling of teachers and a degradation of the profession, should tasks become automated (Selwyn, 2019). Overall, there is a "technological solutionism" (Selwyn, 2019, p. 18) at work in which social problems are said to be solvable by technical means. This obscures the fact that AI and other algorithms are not a neutral tool but are interspersed with social interests and power (Beer, 2017; Williamson, 2015).

How to tame AIED?

These potential problems raise questions with regard to an appropriate reaction. How can we ensure that AIED is used in a responsible way to tackle these challenges? In this regard, three different ways of taming AI can be identified. The first involves an adequate representation of its potential, as well as its shortcomings, and a re-imagining of its envisioned futures. Another possible solution may be to look for regulation and the third option is concerned with strengthening individuals and their capacities.

1. Taming concepts and re-imagining AIED

Against a technological background, a number of authors suggest viewing AI as an element of larger ecosystems. AI is not seen as a neutral tool or a self-sustained powerful force, but is considered as entangled in social worlds, technical infrastructures, and economic as well as political structures. AI is thus "neither *artificial* nor *intelligent*" (Crawford, 2021, p. 9) but a product of (often invisible) human labor, the use of natural resources, and existing classifications. Such a relational view sees AIED as part of educational practices, thus undermining its status as some sort of neutral and objective tool. As with other AI-systems, AIED generates non-transparent results (Perrotta & Selwyn, 2020) and is based on flawed datasets ("broken data") (Pink et al., 2018).

AI and other algorithms form part of larger political and power structures (Bucher, 2018). The use of AIED is thus bound up with policies and strategies pushing the digital transformation of education. Referring to AI in education is, consequently, also a strategy to push educational, political, and economic agendas. As such, AI evokes powerful imaginaries and visions of future education. The future is, however, not determined and AIED is not a deterministic force. Different scenarios are, therefore, not only imaginable but can be enacted by us all. Students can become "smooth users", improved by educational technology or "digital nomads" that individually choose their path from the technological options provided. Alternatively, they could partake in "collective agency" using technology to participate in democratic processes defining the meaning of education (Macgilchrist et al., 2020). It is the responsibility of all the stakeholders involved to determine the course of AIED.

2. Regulating AIED

Closely related to the re-imagining and re-shaping of AI are attempts to regulate its use. AI, in general, is under close scrutiny. The European Commission is, for example, adamant about advocating the advance of these technologies, in order to ensure the economic and societal competitiveness of the European Union (EC, 2021a). At the same time, regulations are proposed dealing with the risks of using AI. This is seen as a particular European approach to AI, so as to create "trustworthy AI" (EC, 2021b, p. 1) and to ensure that AI "works for people" (EC, 2021b, p. 5). Consequently, the "Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment", published by the European Commission, highlights values such as fairness, human agency, and transparency (EC, 2020).

Education, in particular, warrants caution according to the European Commission. The use of AIED systems in education is considered "high-risk", "since they may determine the educational and professional course of a person's life and therefore affect their ability to secure their livelihood" (EC, 2021, p. 26). To ensure that algorithmic decisions regarding educational biographies are made responsibly, transparency and fairness need to be ensured. Along these lines, the European Commission set up an expert group tasked with developing ethical guidelines which deal with the responsible use of AI and data in education and training.[3]

3. Empowering users of AIED

However, besides these proposals and reports, there is no legislation in effect that governs these issues. There is, however, a growing awareness of ethical issues and the fact that users need to be equipped with the appropriate skills to deal with them. Another way of taming AIED is thus to make sure that its users are equipped with adequate competencies and skills.

This is in line with other attempts to delineate digital competencies (e.g., EC, 2016). Data literacy, in particular, is seen as an important skill in the 21st century. In an increasingly datafied world, citizens need to understand how to interpret and analyze data. At the same time, "data infrastructure literacy" (Gray et al., 2018) is required, that is "the ability to account for, intervene around and participate in the wider socio-technical infrastructures through which data is created, stored and analysed" (Gray et al., 2018, p. 1). In the case of education, this would mean that teachers and to some extent students should not only be able to understand the data produced by AI but also *how* they are produced and processed, *by whom*, and *for what* purposes.[4]

Who controls the controls?

This ultimately requires us to rethink the way in which we train and educate teachers. Since their role is bound to change, they not only need to become proficient in the use of AIED but (partially) grasp its modus operandi. If AI "will increasingly become the engine of education, and student data the fuel" (Selwyn et al., 2020, p. 2), then teachers will be the engine drivers handling both the engine and its fuel.

In this regard, teachers are needed more than ever before. AIED can provide educators with powerful tools but they should be tamed with the help of highly trained professionals so that they are used responsibly. They are "humans on the loop" (Mellamphy, 2021) ensuring that algorithmic decision-making is accompanied and supervised by professionals, and their expertise should counter AIED's briefly illustrated shortcomings. Ironically, the more control machines exercise, the more control humans need to be able to exercise. In 1984, Larry Hirschhorn already envisioned that "[I]n cybernetic settings workers must control the controls" (Hirschhorn, 1984, p. 2). Ultimately, AIED would (or rather should) not lead to a loss of control and a deskilling of the pedagogic profession but to a shift in teachers' professional role: from knowledge authorities to knowledge and data infrastructure managers.

References

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J., & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management*, 60, 102387.
https://doi.org/10.1016/j.ijinfomgt.2021.102387

Beer, D. (2017). The social power of algorithms. *Information, Communication & Society, 20*(1), 1–13.
https://doi.org/10.1080/1369118X.2016.1216147

Bucher, T. (2018). *If…then: Algorithmic power and politics*. Oxford University Press.
https://doi.org/10.1093/oso/9780190493028.001.0001

Chang, E. (2019). Beyond workforce preparation: Contested visions of 'twenty-first century' education reform. *Discourse: Studies in the Cultural Politics of Education, 40*(1), 29–45.
https://doi.org/10.1080/01596306.2018.1549702

Crawford, K. (2021). *Atlas of AI.* Yale University Press.
https://doi.org/10.12987/9780300252392

EC. (2016). *DigComp 2.0: The digital competence framework for citizens*. Publications Office of the European Union.
https://data.europa.eu/doi/10.2791/607218

EC. (2020). *The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment*. Publications Office of the European Union.
https://data.europa.eu/doi/10.2759/002360

EC. (2021a). *Annexes to the communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions fostering a European approach to artificial intelligence.* Publications Office of the European Union.

EC. (2021b). *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*. Publications Office of the European Union.

Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society, 5*(2), 1–13.
https://doi.org/10.1177/2053951718786316

Hirschhorn, L. (1984). *Beyond mechanization: Work and technology in a postindustrial age*. MIT Press.

Macgilchrist, F., Allert, H., & Bruch, A. (2020). Students and society in the 2020s: Three future 'histories' of education and technology. *Learning, Media and Technology, 45*(1), 76–89.
https://doi.org/10.1080/17439884.2019.1656235

Mellamphy, N. B. (2021). Humans "in the loop"? Human-centrism, posthumanism, and AI. *Nature + Culture, 16*(1), 11–27.
http://dx.doi.org/10.3167/nc.2020.160102

Perrotta, C., & Selwyn, N. (2020). Deep learning goes to school: Toward a relational understanding of AI in education. *Learning, Media and Technology, 45*(3), 251–269.
https://doi.org/10.1080/17439884.2020.1686017

Pink, S., Ruckenstein, M., Willim, R., & Duque, M. (2018). Broken data: Conceptualising data in an emerging world. *Big Data & Society, 5*(1), 1–13.
https://doi.org/10.1177/2053951717753228

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity.

Selwyn, N., Hillman, T., Eynon, R., Ferreira, G., Knox, J., Macgilchrist, F., & Sancho-Gil, J. M. (2020). What's next for Ed-Tech? Critical hopes and concerns for the 2020s. *Learning, Media and Technology, 45*(1), 1–6.
https://doi.org/10.1080/17439884.2020.1694945

TA-SWISS. (2020). *Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz*. vdf Hochschulverlag.

Williamson, B. (2015). Governing software: Networks, databases and algorithmic power in the digital governance of public education. *Learning, Media and Technology, 40*(1), 83–105.
https://doi.org/10.1080/17439884.2014.924527

Recommended Citation

Röhl, T. (2021). Taming algorithms. *On Education. Journal for Research and Debate, 4*(12). https://doi.org/10.17899/on_ed.2021.12.3

Do you want to comment on this article? Please send your reply to editors@oneducation.net. Replies will be processed like invited contributions. This means they will be assessed according to standard criteria of quality, relevance, and civility. Please make sure to follow editorial policies and formatting guidelines.

tobias röhl

Tobias Röhl is a professor of digital learning and teaching in the Centre for Education and Digital Transformation at the Zurich University of Teacher Education (PH Zürich), Switzerland.

1. https://blogs.lse.ac.uk/impactofsocialsciences/2020/08/26/fk-the-algorithm-what-the-world-can-learn-from-the-uks-a-level-grading-fiasco/
2. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G
3. https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?do=groupDetail.groupDetail&groupID=3774 NB: The author is part of this expert group.
4. This focus on individual competencies, does not mean that the wider ecosystem of AI should not be taken into account. Developers, policy makers and other stakeholders are also accountable for any ethical risks.