



Artificial Intelligence in Education and Ethics

Benedict du Boulay

Contents

Introduction	2
The Early Days of AI in Education	3
Contemporary AI in Education	4
AI and Learner-Facing Tools	4
Are Learner-Facing Tools Effective and Being Used?	5
AI and Teacher-Facing Tools	7
AI and Administrator-Facing Tools	7
Ethical Issues	8
Ethical Issues Around Education in General	8
Ethical Issues of AI in Education	9
Dealing with Ethical Issues in the Design, Implementation, and Deployment of AIED Systems	10
Open Questions and Directions for Future Research	11
Implications for Open, Distance, and Digital Education (ODDE)	12
Conclusion	12
References	13

Abstract

This chapter traces the ethical issues around applying artificial intelligence (AI) in education from the early days of artificial intelligence in education in the 1970s to the current state of this field, including the increasing sophistication of the system interfaces and the rise in data use and misuse. While in the early days most tools were largely learner-facing, now there are tools that are teacher-facing, supporting their management of the classroom, and administrator-facing, assisting in their management of cohorts of students. Learner-facing tools now take into account the affective and motivational aspects of learning as well as the cognitive. The rise of data collection and its associated analytic tools has enabled the development of

B. du Boulay (✉)
University of Sussex, Brighton, UK
e-mail: B.du-Boulay@sussex.ac.uk

dashboards for the dynamic management and reflective understanding of learners, teachers, and administrators. Ethical issues hardly figured in the early days of the field but now they loom large. This is because of the legitimate fears that learners' and teachers' autonomy will be compromised, that learner data will be collected and potentially misappropriated for other purposes, and that AI will introduce extra biases into educational decisions and increase existing inequity and also because of the scary reputation that AI has in general.

Keywords

Artificial intelligence in education · Ethics · Analytics · Dashboards · Learner-facing · Teacher-facing · Administrator-facing

Introduction

Artificial intelligence (AI) was initially applied in education about 50 years ago and only a decade or so after the founding of AI as a research field itself, at a Dartmouth College Workshop, in Hanover, New Hampshire, USA, in 1956 (see, for example, Moor, 2006).

In 1970, Carbonell's paper "AI in CAI: An Artificial-Intelligence Approach to Computer-Assisted Instruction" described a tutor and authoring system named SCHOLAR for geography, based on semantic networks (Carbonell, 1970). This "information structure-oriented (ISO)" tutor separated out its teaching strategy from its knowledge of South American geography in such a way that, in principle, the geography of some other part of the world could be slotted in and the teaching strategy applied to that, or a different teaching strategy applied to the geography of South America. Moreover, because of the explicit representation of its geographic knowledge via semantic networks, the system could reason about its knowledge to draw conclusions that were not explicitly coded in and also answer questions about what it knew. Thus, its "mixed-initiative" teaching strategy could encompass both the system questioning the student, making use of the context and the relevance of its questions, and the student questioning the system, both in very limited English. The system kept track of which bits of the geographical domain had been understood by the student by tagging the relevant parts of the semantic network, thus creating an evolving model of the student's knowledge. This adaptation to the individual learner was one of the factors that distinguished this system from the computer-assisted instructional (CAI) systems that preceded it. The system also exemplified what came to be the standard conceptual architecture of learner-facing artificial intelligence in education (AIEd) systems.

The Early Days of AI in Education

An early collection of AIED papers demonstrated what could already be achieved about a decade later (Sleeman & Brown, 1979). This collection included articles, among others, on systems for computer-based coaching in a gaming scenario (Burton & Brown, 1979), adding tutorial rules to an expert system to enable it to explain and teach the expert system's rules (Clancey, 1979), a knowledge representation to capture the evolving understanding of a learner (Goldstein, 1979), a tutor for elementary programming (Miller, 1979), and a tutoring system for quadratic equations that conducted experiments to evaluate its own teaching performance and then update its own teaching tactics as a result (O'Shea, 1979).

These early papers essentially mapped the conceptual architecture of what are now often called “learner-facing tools,” namely, an explicit model of what is to be taught, an explicit model of how it should be taught, an evolving model of the learner's understanding and skill, and an interface through which the interaction of the learner and the system communicate. Hartley (1973) provided an early definition of this architecture as follows, where (3) and (4) together are the explicit model of teaching, and the interface was not mentioned given its limited scope at that time:

1. A representation of the task
2. A representation of the student and his performance
3. A vocabulary of (teaching) operations
4. A pay-off matrix or set of means-ends guidance rules (Hartley, 1973, p. 424)

The standalone nature of these early systems, their unsophisticated interfaces, and their lack of interest in collecting large amounts of learner data meant that many of the contemporary ethical issues around the use of AIED were not in evidence.

From the start, the general field of AI has had intertwined scientific and engineering aspects (Buchanan, 1988). The scientific aspect of AI *in education* has concerned itself with questions around the nature of human learning and teaching, often with the goal of understanding and then duplicating human expert teaching performance. This aspect has focused largely on learner-facing tools but more recently has expanded into teacher-facing tools. The science has been pursued as a kind of computational psychology for its own sake or as a way to improve educational practice and opportunity in the world. The engineering aspect of applying AIED has exploited a wide range of computational technologies such as Carbonell's semantic networks, mentioned above, and more recently machine learning techniques of various kinds. This aspect of the work has pursued even wider goals that also include the development of educational administrator-facing tools.

This chapter is divided into six sections. Section “[Contemporary AI in Education](#)” gives a brief overview of the current state of the applications of AIED, including subsections on learner-facing tools, teacher-facing tools, and administrator-facing tools. Section “[Ethical Issues](#)” examines the ethical issues that arise from applying AIED, including the ethical issues around educational technology in general, ethical design, and the ethical use and analysis of data. Section “[Open](#)

[Questions and Directions for Future Research](#)” looks at open questions. Section [“Implications for Open, Distance, and Digital Education \(ODDE\)”](#) examines the implications for open, distance, and digital education (ODDE). Section [“Conclusion”](#) offers some brief conclusions.

Contemporary AI in Education

These days the field of AIED has split into three broad overlapping enterprises. The first continues to develop educational tools that focus on *learners* by undertaking various pedagogical roles such as tutoring a set of skills (Koedinger & Aleven, 2016) or assisting concept acquisition (Biswas, Segedy, & Bunchongchit, 2016) or supporting metacognitive awareness and regulation (Azevedo & Aleven, 2013), among others. The second enterprise is the development of assistive tools for teachers (see section [“AI and Teacher-Facing Tools”](#)), and the third enterprise develops tools designed to help educational administrators (see section [“AI and Administrator-Facing Tools”](#)). A useful summary of the applications of AIED for a reader working within ODDE can be found in Kose and Koc (2015).

AI and Learner-Facing Tools

As an example of a tool that focuses on learners, Betty’s Brain is a system designed to help students develop their understanding of the concepts of ecology (Biswas et al., 2016). In this system, the interface is one of the key parts of the system. The student uses the interface to draw a conceptual map consisting of nodes and arrows depicting some of the processes involved in a river ecosystem, such as the absorption of oxygen and the generation of carbon dioxide. The system also provides reading materials from which the student is expected to create the conceptual map. At any time, the student can ask the system to check and test her conceptual map for accuracy and completeness, and the system will offer comments to help her build a better conceptual map. The system is presented in terms of a story where the student is building a conceptual map for an artificial student, Betty, hence Betty’s Brain. The checking and testing is presented as if being set and marked by an artificial teacher, Mr. Davis. Mr. Davis also provides metacognitive hints to the student if she seems to not be paying proper attention to her own learning, such as not making good use of the available reading material.

One of the developments of AIED since the early days has been the focus on learners as human beings with *feelings* and *aspirations* as well as knowledge and skills. This broader focus on the nature of learners and learning has been provoked by our increased understanding of learner motivation (Schunk, Pintrich, & Meece, 2008), mindset (Dweck, 2002), and academic feelings/emotions (Pekrun, 2014), to name but three aspects of human learning. While such an evolution helps to humanize the interaction between systems and learners, it opens up further scope for ethical issues around privacy and around the kinds of data that are collected and

stored. This enlarged focus has involved the development of techniques to try to assess the transient emotional and motivational states of learners in order to boost positive frames of mind, such as engaged concentration, and counter negative states of mind, such as frustration or boredom.

An example of the application of the above is found in a tutor for school mathematics. Arroyo et al. (2014) drew on the work of Dweck (2002) and others to augment an existing tutoring system for mathematics by clustering students' learning behaviors into a small number of profiles in terms of their use of hints, the time they were taking in solving problems, and the number of errors they were making. Each of these profiles was determined by both cognitive and affective/motivational dimensions. For each profile, there were cognitive and affective/motivational actions and feedback by the tutor, such as to set a harder problem (cognitive), praise effort or to de-emphasize the importance of immediate success (affective/motivational).

The suite of language learning tools, Enskill, provides another example of a contemporary interface for learner-facing systems (Johnson, 2019). This is a suite of tools for learning a language, the contextualized use of the correct language register, and for learning how to speak effectively, e.g., making a forceful case for some course of action. The tools use game-based technology to set up an on-screen scenario containing one or more characters with whom the learner speaks and who can reply in speech. The analysis and feedback of the learner's language can be at different levels depending on the context, e.g., pronunciation, grammar, and appropriateness. Moreover, the tools log all interactions with learners, and these link into a mechanism to improve the systems' performance when mistakes or glitches occur ("data-driven development (D3) of learning environments").

A particular outcome of learner analytic aspect of AI in education has been the growth of "dashboards" (Schwendimann et al., 2017). These can be aimed at students to help them reflect on progress, either in the moment or after a lesson or session, or even reflect on the efficacy of the reflection tools themselves (Jivet, Wong, Scheffel, Specht, & Drachsler, 2021). Dashboards aimed at students have grown out of an earlier learner-facing technology named "Open Learner Models" (see, for example, Bull & Kay, 2016).

Are Learner-Facing Tools Effective and Being Used?

There have been at least seven meta-studies and meta-analyses of the effectiveness of learner-facing tools as compared to either a teacher working with a whole class of students or a skilled teacher working with a single student (for a summary, see du Boulay, 2016). The overall message from 182 comparative studies is that learner-facing tools perform better in terms of learning gains compared to a human teacher working with a whole class (effect size = 0.47) but slightly worse than a skilled human tutor working with a single student (effect size = -0.19). In addition to the seven meta-studies, there was a 2-year, large, multistate evaluation of the Cognitive Tutor for Algebra in matched pairs of schools in the USA (Pane, Griffin, McCaffrey,

& Karam, 2014). Each pair consisted of a school that continued to teach algebra in their own fashion and another in which the school also made use of the Cognitive Tutor for Algebra (though without necessarily using it as per the advice of the tutoring system's designers). In the second year of the study, when the teachers had got used to deploying the tutoring system effectively, there was a small comparative learning gain in favor of the schools using the tutoring system (effect size = 0.21).

Despite the positive results for learner-facing tools above, the penetration of artificial intelligence tools of all kinds into schools and colleges has been slow, but with some notable exceptions, such as the Cognitive Tutors in the USA, mentioned above, and now trading under the name Carnegie Learning (Koedinger & Aleven, 2016). More positively, Baker, Smith, and Anissa (2019) say:

Despite minimal attention, AIED tools are already being used in schools and colleges in the UK and around the world – today.

We find learner-facing tools, such as adaptive learning platforms that 'personalise' content based on a child's strengths and weaknesses. We find teacher-facing tools, such as those which automate marking and administration (one government-backed pilot in China sees children in around 60,000 schools having their homework marked by a computer). We find system-facing tools, such as those which analyse data from across multiple schools and colleges to predict which are likely to perform less well in inspections. (p. 5)

According to a systematic review of research on artificial intelligence applications in higher education, the penetration into universities is still patchy with few papers referring either to the ethical dimensions or to learning theory (Zawacki-Richter, Marín, Bond, & Gouverneur, 2019):

The descriptive results show that most of the disciplines involved in AIED papers come from Computer Science and STEM, and that quantitative methods were the most frequently used in empirical studies. The synthesis of results presents four areas of AIED applications in academic support services, and institutional and administrative services: 1. profiling and prediction, 2. assessment and evaluation, 3. adaptive systems and personalisation, and 4. intelligent tutoring systems. The conclusions reflect on the almost lack of critical reflection of challenges and risks of AIED, the weak connection to theoretical pedagogical perspectives, and the need for further exploration of ethical and educational approaches in the application of AIED in higher education. (Zawacki-Richter et al., 2019)

In their editorial to a special issue on AI in university education, that included the paper mentioned above, the editors noted that "there is little evidence at the moment of a major breakthrough in the application of 'modern' AI specifically to teaching and learning, in higher education, with the exception of perhaps learning analytics" (Bates, Cobo, Mariño, & Wheeler, 2020).

AI and Teacher-Facing Tools

Recently, there has been the development of educational tools that focus on teachers to help them either *orchestrate* the use of classroom technology or reflect on that organization. They also (i) help teachers allocate their precious time effectively to those students who need it most and (ii) analyze students' work to determine which are the common issues within a class. We can see this as an evolution of the learner model to encompass both the individuals within a group and the group itself.

For example, the Lumilo system gave the teacher glasses that provided an augmented reality view of her class of students, each working alone with an AIED system (Holstein, McLaren, & Aleven, 2018). There were two kinds of augmentation in this view. The first involved an augmented reality symbol, apparently hovering above each student's head, that indicated their current learning state. These symbols included those for designating the following learner states: idle, (too) rapid attempts, hint abuse or gaming the system, high local error after hints, or unproductive persistence. These symbols were designed to give the teacher information on which to base her decision about which student she should go and help in person. The second augmentation involved an analysis of how the students were doing as a whole to provide a synopsis of problems common to the class. This synopsis was designed to give the teacher information on what might be the focus of her whole class interventions.

AI and Administrator-Facing Tools

The third broad area for AIED has been the rise of analytics applied to data generated in educational contexts at the class or cohort level and aimed at administrator-facing tools. These kinds of analysis explore, for example, the relation of learner engagement to overall success in massive online open courses (MOOCs) (see, for example, Rienties et al., 2016), different patterns of engagement (see, for example, Rizvi, Rienties, Rogaten, & Kizilcec, 2020), and identification of individual and whole class difficulties with course material and the means to rapidly identify and fix any problems and failings in the interactions of a systems with its learners (see, for example, Johnson, 2019).

For example, Peach, Yaliraki, Lefevre, and Barahona (2019) analyzed learners' temporal behavior in online courses at Imperial College Business School and the UK Open University. This data included task completion, timing, and regularity of interactions with the learning system. They mapped individuals' task completion times against the average for all learners and used clustering techniques to create groups that included early birds, on time, low engagers, sporadic outliers, and crammers. They found that poor performers (based on outcome measures) typically evidenced cramming behavior (no surprise there) but good performers were found in all of the time-related groupings, including low engagers and crammers.

In their wide-ranging systematic review, Zawacki-Richter et al. (2019) found a number of papers related to the application of AI in admissions decisions. For

example, Acikkar and Akay (2009) used machine learning techniques to generate a predictive model of whether students would be admitted to university to study physical education and sports based on their “performance in the physical ability test as well as [their] scores in the National Selection and Placement Examination and graduation grade point average (GPA) at high school” (p. 7228). These analyses were undertaken retrospectively and were very accurate (e.g., >90%). The ethical dimension of such predictions comes into sharp focus if such predications are made prospectively when students apply, either as advice to admission tutors or more worryingly as actual decisions with no human in the loop.

Ethical Issues

There have been fears about artificial intelligence from long before the advent of the field (see, for example, *The Golem* (Meyrink, 1915) – a retelling of an ancient tale about animating a living being from a clay statue). However, there were few ethical issues uppermost in the minds of the early creators of student-facing tools using AI. For them, the issues were largely technological and pedagogical, e.g., how to build such systems at all and to determine whether they were effective in educational contexts. These days ethical issues have become much more pressing because of the greater penetration of educational technology (including AI-based technology) into education and training at all levels, the much greater collection of data in educational contexts, and the entry of companies engaged in surveillance capitalism into the educational ecosystem (Williamson, 2018).

Ethical Issues Around Education in General

In most countries, human teachers already operate within an ethical framework. In Scotland, for example, this covers a number of areas, including *doing one's best* for one's students, e.g., by keeping up to date with changes in the curriculum, and *treating students equitably*. It also includes respecting students' confidentiality (see, for example, General Teaching Council Scotland, 2012).

The rise of educational technology of all kinds, whether involving AI or not, and its creation of logs of interactions, has produced a huge amount of student data at all levels in education from primary (elementary) schools to universities. Teachers' ethical guidelines, such as those above, need to encompass these extra sources of data. There are many unanswered questions about who owns this data, who has access to it, how long it will be kept, and so on. The European Framework on General Data Protection Regulation (GDPR) provides guidance on managing all kinds of personal data (Li, Yu, & He, 2019). However, there are still issues for students around understanding what data about them counts as “personal” (Marković, Debeljak, & Kadoić, 2019), as well as around their degree of ownership and rights over educational log data.

Ethical Issues of AI in Education

Involving AI into educational technology must also be required to *do its best and treat students equitably*. For learner-facing tools, one should expect that designers of the educational technology will ensure that the technology will do the best that is possible in the circumstances, whether it is teaching, tutoring, mentoring, or counseling students. One should also expect that the technology treats students in an equitable fashion and does not favor one student over another either inadvertently or deliberately.

Learner-Facing Tools

How might it ever be the case that technology treats students inequitably, we may ask? Many learner-facing systems select what they think is the next most useful learning task, e.g., the next problem to solve, for a particular student with a particular educational history. A faulty design-level categorization of learners into groups, e.g., by gender, perceived prior attainment, motivation, or self-regulated learning capability, might lead to a student being presented with inappropriate or much tougher tasks than they can cope with or indeed much easier tasks than they can learn from. Of course, this kind of bias can also happen with human teachers where their low expectations of some students can become self-fulfilling prophecies. But just because human teachers can, on occasion, be biased does not mean we should turn a blind eye to the potential biases in AI-based educational technology.

Teacher-Facing Tools

Similar considerations apply to teacher-facing tools. The Lumilo orchestration system we described above flagged up students who were doing OK or who were experiencing different kinds of difficulty. This aimed to enable the teacher to make choices about who to help. Clearly, this is an ethically charged decision. Should the teacher prioritize those who are in most difficulty or spread her effort more evenly across the whole class? That is a human dilemma. But the orchestration system had better get its diagnostics correct about who it thinks is doing OK and who it thinks needs help. Even without the use of AI, systems for helping the teacher manage a classroom can have unexpected negative effects on the students. In a study of ClassDojo, used by teachers to record student behaviors, Lu, Marcu, Ackerman, and Dillahun (2021) noted that:

In particular, the use of ClassDojo runs the risk of measuring, codifying, and simplifying the nuanced psycho-social factors that drive children's behavior and performance, thereby serving as a "Band-Aid" for deeper issues. We discuss how this process could perpetuate existing inequality and bias in education. (Lu et al., 2021)

Administrator-Facing Tools

For a discussion of the wide uses of AI in universities, see Zeide (2019). Administrator-facing tools are sometimes used to make predictions about which students seem to be doing broadly OK and which are showing evidence of failing the

course or dropping out. These kinds of judgment are often based on learning analytics using AI methodologies. The issue here is the consequences of false negatives and false positives emerging from an inadequate data analysis. For example, missing signs that a student is really struggling may mean that no human is alerted to provide help. Labeling a student as struggling who is doing OK may also have repercussion down the line, rather like an incorrect entry in a credit rating. For an interesting example of the artifacts that can occur in analyzing cohort data, see Alexandron, Yoo, Ruipérez-Valiente, Lee, and Pritchard (2019). They showed that sometimes students using a MOOC set up two accounts so that they could game the system. One account (in a fake name) would be used to get lots of help from the system to find the right answers, while the other account (in the student's real name) would be used to answer all the questions quickly and correctly.

Using predictions to drive admissions of students to schools or colleges (Acikkar & Akay, 2009) or to predict grades when exams could not be taken because of COVID are fraught with ethical issues. The recent creation and then abandonment of an algorithm to predict UK student grades for entry to university is a salutary reminder about both potential AI biases and the potential human teacher biases the algorithm was intended to mitigate (see, for example, Hao, 2020).

Dealing with Ethical Issues in the Design, Implementation, and Deployment of AIED Systems

Many different systems are now designed that include AI elements, from smartphone apps to big bank data systems. There has been increasing concern about the ethical questions that arise in the design, implementation, and deployment of such systems, with the EU proposing legislation to manage the situation (European Commission, 2020). Many different frameworks have been proposed to manage the development of such systems. A useful summary of such frameworks can be found in Floridi and Cows (2019). They developed a framework from bioethics, under the general headings of beneficence, non-maleficence, autonomy, and justice, to also include "explicability." Most systems should work under the control of (or at least in tandem with) humans, so it is important that the system employing AI is able to offer an explanation or justification for exactly why it is suggesting a decision, a course of action, an outcome, or whatever, in order that the human can weigh up the degree to which he or she should agree with the machine. Particularly in education, autonomy and explicability must play a central role.

The issue of collecting, analyzing, and managing learner data has become more pressing for many reasons, including (i) greater general awareness of data privacy issues, (ii) the sheer quantity of learner data being collected, (iii) the increased use of AI and other methodologies for finding patterns in that data and drawing inferences from them, and (iv) the use of learner (and thus user) data for commercial purposes which have nothing to do with education (Williamson, 2018).

For example, Williamson (2018) warns about “Big Tech” companies moving into the field of education, typically with learner-facing tools, so that they can harvest the learner data for commercial purposes:

Startup schools are analysed as prototype educational institutions that originate in the culture, discourse and ideals of Silicon Valley venture capital and startup culture, and that are intended to relocate its practices to the whole social, technical, political and economic infrastructure of schooling. These new schools are being designed as scalable technical platforms; funded by commercial “venture philanthropy” sources; and staffed and managed by executives and engineers from some of Silicon Valley’s most successful startups and web companies. Together, they constitute a powerful shared “algorithmic imaginary” that seeks to “disrupt” public schooling through the technocratic expertise of Silicon Valley venture philanthropists. (Williamson, 2018, p. 218)

Researchers within AI in education are starting to be aware of these ethical issues, even though Zawacki-Richter et al. (2019) found only two papers in their systematic review of the applications of AI in universities that dealt with ethical issues. So, for example, we see both the emergence of general design frameworks, such as that of Floridi and Cowls (2019) above, for including AI in software products, and those aimed specifically at the development of AI applications in education (see, for example, Drachsler & Greller, 2016) and, most notably, the creation of an Institute of Ethical AI in Education which has set out guidelines particularly for teachers in their use of applications of AI (Seldon, Lakhani, & Luckin, 2021).

Open Questions and Directions for Future Research

From an ethical point of view, the big issue is how we can ensure that learners acquire more control over the data that is generated when they interact with educational technology and are protected from the misuse of their data by others. This section identifies some open questions and directions of research in the science and engineering of applications of AIED for each of the categories of system mentioned above, namely, learner-facing, teacher-facing, and administrator-facing.

Given the increasing interest in gathering and using affective data about learners to improve the adaptivity of learner-facing tools, two *scientific* questions are (i) what might be the most useful affective categories on which to develop an affective pedagogy and (ii) what kinds of pedagogic rules should be used to maximize the chance of fruitful learning, given the sequence of the learner’s cognitive and affective states so far. For example, should hope, dismay, and pride also play a role as well as confusion, frustration, and engaged concentration, and how should they be “managed”?

For systems aimed at teachers, an *engineering* question is: How best to manage and support the division of labor between the human teacher(s) and the system, given the manifest complexity and dynamic nature of most classrooms full of learners? For example, how should a tool used for dynamic management differ from one used for reflective practice?

For systems aimed at understanding cohorts, an *engineering* question is: How best can learning management systems be developed to measure and potentially answer *academic* questions about learning rather than *administrative* ones? For example, did students on this course show strong evidence of improvements in their self-regulated learning capability?

Implications for Open, Distance, and Digital Education (ODDE)

There are three main implications for ODDE. The first is that one of the oldest technologies for distance learning, the textbook, has been enhanced by the application of AI, either through adapting the content or the route through that content to the reader (see, for example, Thaker, Huang, Brusilovsky, & He, 2018). The second implication is that online, distance, and digital systems have increasingly incorporated elements of AI in order to make such systems smarter and more responsive to the needs of learners and teachers (see, for example, Kose, 2015; UNESCO, 2021). The third implication is that the developers and deployers of ODDE systems are already taking an ethical stance on how the systems are designed and built, how they are used in practice, and how their data is collected, stored, and analyzed (Prinsloo & Slade, 2016). For example, with particular respect to ODDE, Sharma, Kawachi, and Bozkurt (2019) state:

First, there should be some control mechanisms that should be put into place to ensure transparency in collection, use and dissemination of the AI data. Second, we need to develop ethical codes and standards proactively so that we truly benefit from AI in education without harming anything; not only humans but any entity. Third, we should ensure learners' privacy and protect them for any potential harm. Next, we must raise awareness about the AI so that individuals can protect themselves and take a critical position when needed. (p. 2)

Conclusion

In order to give a longitudinal view of AIEd and ethics, this chapter has sketched the early days of learner-facing system development in the 1970s as well as provided some examples of much more recent systems. While the early systems were mostly learner-facing, contemporary applications of AI now also include teacher-facing and administrator-facing tools and are used both locally and via online, distance, and digital technologies.

The interface is one area where there have been big changes. One of the earliest systems had an interface that involved the learner typing in answers (and indeed questions) in stilted English, whereas contemporary learner-facing tools can show lifelike pedagogical agents with whom learners have a spoken dialogue in everyday English. Moreover, tools for other kinds of user make use of complex interactive dashboards.

In the early days, learner-facing tools were largely designed to work as tutors with a single learner. These days some tools are still designed to work with a single learner, though they can now adapt to the learner's affective and motivational state as well as to what the learner knows and understands. Other tools can work with more than a single learner (see, for example, Walker, Rummel, & Koedinger, 2009), and others again work with teachers rather than learners to assist them in the complex task of managing a class full of students and allocating their limited time in the most effective way.

The creation of log data from educational systems and the use of data mining and other analytic techniques have given rise to the thriving field of learner analytics. This in turn has enabled the creation of dashboards for learners, teachers, and administrators to interrogate data at varying levels of granularity.

Ethics has played a strong role in education for many years, most obviously via the codes of professional practice that teachers are expected to act within. In the early days of AI, ethical issues around education tool design and deployment were not uppermost in the minds of designers: simply making the systems work effectively was the main goal. Nowadays, ethics is very much in people's minds whether they be system designers, teachers, parents, administrators, or indeed learners, but there is still a long way to go to make educational technology a place of trust and safety.

AI has a mixed reputation. On the one hand, it is so ubiquitous that we hardly notice it, e.g., interacting with a chatbot on a website or having one's camera optimize a photograph. On the other hand, there are scary stories about AI taking over the world, or just as scary reports about biased decisions that might affect one's well-being (e.g., refusal of a mortgage or a job) or one's life (e.g., a diagnostic system generating a false-positive or false-negative report about a tumor). Within education, there are issues about the ways that analytics may produce biased results or that companies using AI enter education not with learner's best interest at heart but as a way to hoover up their data for commercial purposes. To counter these issues, various codes of ethics have been developed that cover all aspects of the design and deployment of AI-based educational technology at both international and more local levels.

Acknowledgments I thank Mutlu Cukurova, Carmel Kent and Anissa Moeini and the reviewers for helpful suggestions on a draft of this paper.

References

- Acikkar, M., & Akay, M. F. (2009). Support vector machines for predicting the admission decision of a candidate to the School of Physical Education and Sports at Cukurova University. *Expert Systems with Applications*, 36, 7228–7233. <https://doi.org/10.1016/j.eswa.2008.09.007>
- Alexandron, G., Yoo, L. Y., Ruipérez-Valiente, J. A., Lee, S., & Pritchard, D. E. (2019). Are MOOC learning analytics results trustworthy? With fake learners, they might not be! *International Journal of Artificial Intelligence in Education*, 29, 484–506. <https://doi.org/10.1007/s40593-019-00192-0>

- Arroyo, I., Woolf, B. P., Burleson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education*, 24, 387–426.
- Azevedo, R., & Aleven, V. (Eds.). (2013). *International handbook of metacognition and learning technologies*. New York: Springer.
- Baker, T., Smith, L., & Anissa, N. (2019). Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges. Retrieved from https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf
- Bates, T., Cobo, C., Mariño, O., & Wheeler, S. (2020). Can artificial intelligence transform higher education? *International Journal of Educational Technology in Higher Education*, 17(1), 42. <https://doi.org/10.1186/s41239-020-00218-x>
- Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From design to implementation to practice a learning by teaching system: Betty's Brain. *International Journal of Artificial Intelligence in Education*, 26, 350–364. <https://doi.org/10.1007/s40593-015-0057-9>
- Buchanan, B. G. (1988). Artificial intelligence as an experimental science. In J. H. Fetzer (Ed.), *Aspects of artificial intelligence* (pp. 209–250). Dordrecht: Kluwer Academic Publishers.
- Bull, S., & Kay, J. (2016). SMILI[©]: A framework for interfaces to learning data in open learner models, learning analytics and related fields. *International Journal of Artificial Intelligence in Education*, 26, 293–331. <https://doi.org/10.1007/s40593-015-0090-8>
- Burton, R. R., & Brown, J. S. (1979). An investigation of computer coaching for informal learning activities. *International Journal of Man-Machine Studies*, 11, 5–24. [https://doi.org/10.1016/S0020-7373\(79\)80003-6](https://doi.org/10.1016/S0020-7373(79)80003-6)
- Carbonell, J. R. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE Transactions on Man-Machine Systems*, 11, 190–202. <https://doi.org/10.1109/TMMS.1970.299942>
- Clancey, W. J. (1979). Tutoring rules for guiding a case method dialogue. *International Journal of Man-Machine Studies*, 11, 25–50. [https://doi.org/10.1016/S0020-7373\(79\)80004-8](https://doi.org/10.1016/S0020-7373(79)80004-8)
- Drachler, H., & Greller, W. (2016). Privacy and analytics: It's a DELICATE issue a checklist for trusted learning analytics. In S. Dawson, H. Drachler, & C. P. Rosé (Eds.), *Enhancing impact: Convergence of communities for grounding, implementation, and validation* (pp. 89–98). Edinburgh: ACM. <https://doi.org/10.1145/2883851.2883893>
- du Boulay, B. (2016). Artificial intelligence as an effective classroom assistant. *IEEE Intelligent Systems*, 31(6), 76–81. <https://doi.org/10.1109/MIS.2016.93>
- Dweck, C. S. (2002). Beliefs that make smart people dumb. In R. J. Sternberg (Ed.), *Why smart people can be so stupid* (pp. 24–41). New Haven: Yale University Press.
- European Commission. (2020). White Paper on Artificial Intelligence: A European approach to excellence and trust (COM(2020) 65 final). Retrieved from https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en
- Floridi, L., & Cows, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1), 1–13. <https://doi.org/10.1162/99608f92.8cd550d1>
- General Teaching Council Scotland. (2012). Code of Professionalism and Conduct. Retrieved from <https://www.gtcsc.org.uk/regulation/copac.aspx>
- Goldstein, I. P. (1979). The genetic graph: A representation for the evolution of procedural knowledge. *International Journal of Man-Machine Studies*, 11, 51–78. [https://doi.org/10.1016/S0020-7373\(79\)80005-X](https://doi.org/10.1016/S0020-7373(79)80005-X)
- Hao, K. (2020). The UK exam debacle reminds us that algorithms can't fix broken systems. *The MIT Technology Review*. Retrieved from <https://www.technologyreview.com/2020/08/20/1007502/uk-exam-algorithm-cant-fix-broken-system/>
- Hartley, J. R. (1973). The design and evaluation of an adaptive teaching system. *International Journal of Man-Machine Studies*, 5, 421–436. [https://doi.org/10.1016/S0020-7373\(73\)80029-X](https://doi.org/10.1016/S0020-7373(73)80029-X)
- Holstein, K., McLaren, B. M., & Aleven, V. (2018). Student learning benefits of a mixed-reality teacher awareness tool in ai-enhanced classrooms. In C. P. Rosé, R. Martínez-Maldonado, H. U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, B. McLaren, & B. du Boulay (Eds.),

- Artificial intelligence in education: 19th international conference, AIED 2018, London, UK, June 27–30, 2018 proceedings, part I* (pp. 154–168). Cham: Springer.
- Jivet, I., Wong, J., Scheffel, M., Specht, M., & Drachsler, H. (2021). Quantum of choice: How learners' feedback monitoring decisions, goals and self-regulated learning skills are related. In M. Scheffel, N. Dowell, S. Joksimovic, & G. Siemens (Eds.), *The impact we make: The contributions of learning analytics to learning* (pp. 416–427). Edinburgh: ACM. <https://doi.org/10.1145/3448139.3448179>
- Johnson, W. L. (2019). Data-driven development and evaluation of Enskill English. *International Journal of Artificial Intelligence in Education*, 29, 425–457. <https://doi.org/10.1007/s40593-019-00182-2>
- Koedinger, K. R., & Aleven, V. (2016). An interview reflection on “Intelligent Tutoring Goes to School in the Big City”. *International Journal of Artificial Intelligence in Education*, 16, 13–24. <https://doi.org/10.1007/s40593-015-0082-8>
- Kose, U. (2015). On the intersection of artificial intelligence and distance education. In U. Kose & D. Koc (Eds.), *Artificial intelligence applications in distance education* (pp. 1–11). Hershey: IGI Global.
- Kose, U., & Koc, D. (2015). *Artificial intelligence applications in distance education*. Hershey: IGI Global.
- Li, H., Yu, L., & He, W. (2019). The impact of GDPR on global technology development. *Journal of Global Information Technology Management*, 22, 1–6. <https://doi.org/10.1080/1097198x.2019.1569186>
- Lu, A. J., Marcu, G., Ackerman, M. S., & Dillahunt, T. R. (2021). *Coding bias in the use of behavior management technologies: Uncovering socio-technical consequences of data-driven surveillance in classrooms*. Paper presented at the DIS'21: Conference on Designing Interactive Systems, Virtual Event, USA. Retrieved from https://static1.squarespace.com/static/5ebb1d874617b44f913c6d4b/t/609afa7f6ca7b40f39e55106/1620769435258/lu_dis21.pdf
- Marković, M. G., Debeljak, S., & Kadoić, N. (2019). Preparing students for the era of the General Data Protection Regulation (GDPR). *TEM Journal: Technology, Education, Management, Informatics*, 8, 150–156. <https://doi.org/10.18421/TEM81-21>
- Meyrink, G. (1915). *Dr Golem*. Leipzig: Kurt Wolff.
- Miller, M. L. (1979). A structured planning and debugging environment for elementary programming. *International Journal of Man-Machine Studies*, 11, 79–95. [https://doi.org/10.1016/S0020-7373\(79\)80006-1](https://doi.org/10.1016/S0020-7373(79)80006-1)
- Moor, J. (2006). The Dartmouth College Artificial Intelligence conference: The next fifty years. *AI Magazine*, 27(4), 87–91. <https://doi.org/10.1609/aimag.v27i4.1911>
- O'Shea, T. (1979). A self-improving quadratic tutor. *International Journal of Man-Machine Studies*, 11, 97–124. [https://doi.org/10.1016/S0020-7373\(79\)80007-3](https://doi.org/10.1016/S0020-7373(79)80007-3)
- Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at scale. *Educational Evaluation and Policy Analysis*, 36, 127–144. <https://doi.org/10.3102/0162373713507480>
- Peach, R. L., Yaliraki, S. N., Lefevre, D., & Barahona, M. (2019). Data-driven unsupervised clustering of online learner behaviour. *npj Science of Learning*, 4, 14. <https://doi.org/10.1038/s41539-019-0054-0>
- Pekrun, R. (2014). Emotions and learning. Retrieved from <http://www.ibe.unesco.org/en/document/emotions-and-learning-educational-practices-24>
- Prinsloo, P., & Slade, S. (2016). Big data, higher education and learning analytics: Beyond justice, towards an ethics of care. In K. D. Ben (Ed.), *Big data and learning analytics in higher education: Current theory and practice* (pp. 109–124). Cham: Springer.
- Rienties, B., Boroowa, A., Cross, S., Farrington-Flint, L., Herodotou, C., Prescott, L., ... Woodthorpe, J. (2016). Reviewing three case-studies of learning analytics interventions at the Open University UK. In S. Dawson, H. Drachsler, & C. P. Rosé (Eds.), *Enhancing impact: Convergence of communities for grounding, implementation, and validation* (pp. 534–535). Edinburgh: ACM. <https://doi.org/10.1145/2883851.2883886>

- Rizvi, S., Rienties, B., Rogaten, J., & Kizilcec, R. F. (2020). Investigating variation in learning processes in a FutureLearn MOOC. *Journal of Computing in Higher Education*, 32, 162–181. <https://doi.org/10.1007/s12528-019-09231-0>
- Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2008). *Motivation in education: Theory, research and applications* (3rd ed.). Upper Saddle River: Pearson/Merrill Prentice Hall.
- Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., . . . Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, 10, 30–41. <https://doi.org/10.1109/tlt.2016.2599522>.
- Seldon, A., Lakhani, P., & Luckin, R. (2021). The ethical framework for AI in education. Retrieved from <https://www.buckingham.ac.uk/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-The-Ethical-Framework-for-AI-in-Education.pdf>
- Sharma, R. C., Kawachi, P., & Bozkurt, A. (2019). The landscape of artificial intelligence in open, online and distance education: Promises and concerns. *Asian Journal of Distance Education*, 14, 1–2. Retrieved from <http://www.asianjde.com/ojs/index.php/AsianJDE/article/view/432>
- Sleeman, D. H., & Brown, J. S. (1979). Editorial: Intelligent tutoring systems. *International Journal of Man-Machine Studies*, 11, 1–3. [https://doi.org/10.1016/S0020-7373\(79\)80002-4](https://doi.org/10.1016/S0020-7373(79)80002-4)
- Thaker, K., Huang, Y., Brusilovsky, P., & He, D. (2018). *Dynamic knowledge modeling with heterogeneous activities for adaptive textbooks*. Paper presented at the 11th International conference Educational Data Mining (EDM 2018), Buffalo.
- UNESCO. (2021). The Open University of China awarded UNESCO Prize for its use of AI to empower rural learners. Retrieved from <https://en.unesco.org/news/open-university-china-awarded-unesco-prize-its-use-ai-empower-rural-learners>
- Walker, E., Rummel, N., & Koedinger, K. R. (2009). Integrating collaboration and intelligent tutoring data in the evaluation of a reciprocal peer tutoring environment. *Research and Practice in Technology Enhanced Learning*, 4, 221–251. <https://doi.org/10.1142/S179320680900074X>
- Williamson, B. (2018). Silicon startup schools: Technocracy, algorithmic imaginaries and venture philanthropy in corporate education reform. *Critical Studies in Education*, 59, 218–236. <https://doi.org/10.1080/17508487.2016.1186710>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1). <https://doi.org/10.1186/s41239-019-0171-0>
- Zeide, E. (2019). Artificial intelligence in higher education: Applications, promise and perils, and ethical questions. *Educause Review*, 31–39. Retrieved from <https://er.educause.edu/-/media/files/articles/2019/8/er193104.pdf>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

