



LONDON REVIEW OF EDUCATION

e-ISSN: 1474-8479

Journal homepage:

<https://www.uclpress.co.uk/pages/london-review-of-education>

Artificial intelligence in schools: Towards a democratic future

Sandra Leaton Gray 

How to cite this article

Leaton Gray, S. (2020) 'Artificial intelligence in schools: Towards a democratic future'. *London Review of Education*, 18 (2): 163–177. <https://doi.org/10.14324/LRE.18.2.02>

Submission date: 24 September 2019

Acceptance date: 28 April 2020

Publication date: 21 July 2020

Peer review

This article has been peer-reviewed through the journal's standard double-blind peer review, where both the reviewers and authors are anonymized during review.

Copyright

© 2020 Leaton Gray. This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY) 4.0 <https://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Open access

London Review of Education is a peer-reviewed Open Access journal.

Artificial intelligence in schools: Towards a democratic future

Sandra Leaton Gray* – *UCL Institute of Education, UK*

Abstract

The introduction of artificial intelligence in education (AIED) is likely to have a profound impact on the lives of children and young people. This article explores the different types of artificial intelligence (AI) systems in common use in education, their social context and their relationship with the growth of commercial knowledge monopolies. This in turn is used to highlight data privacy rights issues for children and young people, as defined by the 2018 General Data Protection Regulations (GDPR). The article concludes that achieving a balance between fairness, individual pedagogic rights (Bernstein, 2000), data privacy rights and effective use of data is a difficult challenge, and one not easily supported by current regulation. The article proposes an alternative, more democratically aware basis for artificial intelligence use in schools.

Keywords: artificial intelligence, algorithms, data privacy, General Data Protection Regulations (GDPR), Bernstein, Homans

In 2018, the new head of Artificial Intelligence (AI) at the digital education company Pearson was quoted in a company press release as saying:

Unlike other sectors, education is yet to fully realize the benefits of digital and advanced AI techniques, and there are great opportunities to improve learning outcomes and to enable better teaching ... Pearson is committed to transforming the learning experience and *becoming the digital winner* in education. (Pearson, 2018, emphasis added)

This quotation raises a very important and difficult question. In the apparent commercial push for 'digital winners', how do we ensure appropriate levels of democratic accountability, trust and fairness when it comes to introducing artificially intelligent systems into schools?

To answer this question, we need to start by exploring the ways artificial intelligence (AI) is being used, as well as the consequences of adopting different applications for artificial intelligence in education (AIED). It is important to consider AI in this way because no technology can ever be regarded as truly neutral. Indeed, each technological development brings significant repercussions for the political, economic and cultural aspects of society (Feenberg, 1999; Clarke *et al.*, 2007). The need for mutual trust among key stakeholders involved in the education process can, and does, get lost in technological translation. Here we should take the term 'trust' to mean operating technological systems reliably, transparently and with all of our best interests at the centre, a process boyd and Crawford (2012) describe as 'empowerment'.

A concern for fairness in technological processes reflects that of earlier thinkers such as Bernstein (2000) and his heuristic of pedagogic rights, as well as Homans's (1958) understanding of the nature of social exchange, both of which are elaborated later in this article. Bernstein (2000) and Homans (1958) argued in different ways that fairness is essential to effective social relations. When we think about AIED, therefore, it is crucial to look beyond adopting new technologies or processes purely on grounds of efficiency or intellectual augmentation, such as when we are monitoring students, tracking their development, offering new avenues for learning or for governance purposes (Eynon, 2013; Kirkman, 2014; Polonetsky and Tene, 2014; Sclater and Bailey, 2018; Selwyn, 2011a; Williamson, 2015). We need to see how they can contribute to a healthy democracy, nurturing a sense of agency among users, rather than simply mirroring narrow commercial interests, as has happened with the adoption of surveillance technologies such as biometrics in schools.

Even though there is clearly significant work to be done in tailoring the technically and ethically complex field of AI to fit a heterogeneous schooling system, it also represents a great moment of opportunity. For the first time, we are finding that machines are in a position to influence learning and teaching in ways that are far removed from the human capability of their operators, by ceasing to rely on a hypothesis-based method of deduction (Anderson, 2008; Lawn, 2013; Eynon, 2013). This change, which has been called the 'datafication' of education, is much more significant than the early introduction of computing systems. (See Jarke and Breiter (2019) for a wider mapping of the current use of the term 'datafication'.) The reason for this is that the processes they use are starting to surpass the ability of human operators to replicate their functions manually, even in a lengthy manner, as technological systems start to interrogate and teach each other, rather than simply respond to one set of instructions.

This represents a scaling up of data collection and analysis concerning children and young people that is unprecedented and, to a large extent, relatively uncontrolled, with the commercial sector playing a key role in the global massification of provision. The scaling up has associated impacts on privacy (Har Carmel, 2016; Pardo and Siemens, 2014; Selwyn, 2015; Tudor, 2015; Williamson, 2017a). Implications for child-teacher relations and the cohesion of national education systems in the future are consequently profound, and, along with examining conceptions of democratic accountability, trust and fairness, this is a key theoretical focus of this article.

To this end, the first section of the article will provide a broadly synoptic description of different developments in the field of AIED. Together they act as an explanatory text, representing what Mackenzie (2017) describes as an 'accumulation' of techniques with complex and interconnected histories. They are presented here as a means of introducing readers to some of the key roles that AI increasingly plays within contemporary schooling systems, before we move on to explore the social and democratic consequences of their adoption at scale.

AI technologies and some potential applications to education

AI as a concept is understood to have originated in the work of Alan Turing (1950), and the term itself was first used by MIT Professor John McCarthy *et al.* (1955). We may think we understand what AI means, but it is a somewhat inconclusive term, represents a vast array of technological innovations and activities and is perhaps better suited to describing goal-based behaviours rather than the technical means

Table 1: Summary of common terms for current AI technologies and some applications to education

	Advantages	Disadvantages	Term overlaps most closely with	Applications in education
Predictive analysis	Can assess probability of successful outcomes	Can result in adverse selection disadvantaging minority groups	Machine learning, deep learning	Student selection
Deep learning	Recognizes objects, descriptions and people	Some inaccurate and/or discriminatory interpretations of images, objects and texts	Predictive analysis, neural networks, machine learning	Student surveillance, creative curriculum work
Machine learning	Mapping trends and patterns in large data sets	Can identify trends and patterns of little practical use	Predictive analysis, deep learning	School inspection
Neural networks	Identifies patterns and behaviours	Potential for invasions of data privacy	Deep learning, social robotics	School discipline and student monitoring
Expert systems	Knowledge diagnosis and remediation among students	Can result in poor social or cultural fit	Social robotics	Supplementary teaching and student support
Social robotics	Consistent output	Can result in user alienation	Neural networks, deep learning	Personalized learning

of achieving them (Mackenzie, 2017; Russell and Norvig, 2016). As such, the term embraces several different and frequently significantly overlapping concepts, including those described below. A fuller account, beyond the scope of this article, can be found in Luckin (2018). This section acts as a technical and conceptual map to indicate the wider consequences as AIED is rolled out across the sector. In addition, each technique is given an education-related label, in order to bring to life (for the technical lay reader) its potential institutional function. This is summarized in Table 1.

Predictive analysis: 'The digital school bursar'

The very general term 'predictive analysis' dates back to the Good–Turing probability estimator, used initially by Turing to decode German messages in the Second World War, and later refined by Good (1966) to allow for incomplete data sets. It is used to predict future events through the deployment of probabilistic statistical calculations. In a contemporary context, it is frequently used to predict 'likes' or recommendations for online purchases. It is also used for language processing (including autocomplete functions) and for things such as crime and health trend prediction (Abbott, 2014). It can be used in education to assess things such as the likelihood of young people from different backgrounds being successful on different types of academic programme, or to identify students at risk of dropping out (Ong, 2016). This seems a worthy aim, but there are serious democratic considerations when this technique is used, as it may inadvertently screen out minority groups of individuals repeatedly from higher-

level opportunities. It can do this through algorithmic bias and adverse selection, where deprived or vulnerable groups are deliberately rejected from a group, perhaps through automated decision making (Cossins, 2018). An example of this might be the Intelligent Zoning Engine (IZE), which has been used since 2017 to determine optimal catchment areas for the Berlin district of Tempelhof-Schöneberg through calculation of travel time/distance, even though this may inadvertently entrench disadvantage through a form of inadvertent ghettoization of deprived students in particular schools (Algorithm Watch, n.d.). Another example might be the Open University's use of the OU Analyse tracking system (Herodotou *et al.*, 2019), which identifies students requiring tutorial intervention, but that may be limited in scope for wider use given the predominantly mature, part-time university population that the tool has been trained upon. If systems such as these are to be used in a different context, they will need completely retraining on a new data set, as reliability can only ever be population-specific. The dangers of training analytical tools on atypical populations is widely recognized as existing in other fields (see, for example, Goldhaber-Fiebert and Prince, 2019, which discusses accuracy problems in screening in child protection cases).

Deep learning: 'The digital art/music/writing room'

Deep learning uses a type of algorithm that loops repeatedly to calculate increasingly higher-level features. For example, in a facial recognition system used to identify students in a school, the raw data may consist of pixels; the first representational layer may encode some edges; the second representational layer may identify the edges of a face more precisely; the third may encode individual facial features; and the fourth may identify the image as a face. Finally, a system may be able to create an image of a face successfully on its own. Deep learning in the form of facial recognition has been shown to work in some circumstances, but with significant algorithmic bias towards white men, for example (Buolamwini and Gebru, 2018). In a more general sense, deep learning has been used to create rudimentary music and artworks by viewing thousands of examples and extrapolating from these typical aspects of production and representation (Huang and Wu, 2016; Elgammal, 2019). A further affordance has been the creation of works of literature via deep learning using natural language-processing techniques (Hornigold, 2018). In the context of education, this approach could be used for surveillance, such as identifying which people in a crowd are wearing a particular type of school uniform. More interestingly, it may also have curriculum potential through encouraging students to develop artistic works through new forms of analysis and interaction with computers. The term also embraces, to some extent, the concept of spatial augmented reality, in which interactive rendering algorithms are used to create immersive displays. An example of this in an educational context is the use of FUTUREGYM by non-verbal disabled children. This is a system of games in which virtual and real objects coexist, used as a method for encouraging communication (Takahashi *et al.*, 2018).

Machine learning: 'The digital inspector'

Machine learning is a term developed by Samuel (1959) in relation to the development of a computer programme designed to play an online version of the board game checkers (draughts). It refers to the use of algorithms and statistical models to perform tasks without specific instructions, and with little if any human intervention. An example of this is in school inspection, where national attainment trends can be mapped to

identify policy outcomes as well as local and regional anomalies, identifying schools requiring human inspection (Ofsted, 2018). This can be helpful, but it is not entirely unproblematic. While a machine learning system may find data patterns, these may be a subsequence rather than a consequence of human action, as with any statistical analysis. For example, a cluster of students experiencing lower attainment one year may be a coincidence (to do with local weather conditions, or an epidemic of some kind, for example) and have little to do with any school-related provision. This may generate false positives for a school inspection service, triggering inspections where they are not needed (BBC News, 2017; Reynolds, 2017). This is because in such a model, there is a risk that data are seen as neutral and their intrinsic value and subsequent analysis are not questioned.

Neural networks: 'The digital prefect'

The term 'neural networks' refers to a system that is intended to emulate a biological system such as the human brain. There are a number of 'nodes', or artificial neurons, which are taught to recognize whether something is true or false via a system of algorithms. A system can then train independently to identify similar examples more frequently and reliably. For example, it would theoretically be possible to teach a system how to recognize individual students in crowds or in social media images, as in the Chinese social credit system, although in practice a system would need many examples in order to do this sufficiently well, so the possibility may not be realistic (Chorzempa *et al.*, 2018) and scores could potentially be 'gamed' by individuals (Lehr, 2019). There may also be significant privacy violations (Hess, 2019). Another common use is inferring attention or emotion from video streams of classroom activities. This makes presumptions about the social and cultural context of body posture (known as emotion analytics), which may not be relevant or accurate. Systems such as these have been heavily criticized for encouraging behavioural performativity (Manolev *et al.*, 2019; Williamson, 2017b). There have also been recent attempts to use neural networks to build a 'school engagement scale' to draw on data derived from school, family and demographic variables, and aimed at reducing student dropout rates (Turhan *et al.*, 2016). Again, this offers superficial promise, but comes at the risk of discrimination and inaccuracy. One question that may need to be asked is whether different categories are being conflated, for example, ethnicity and social class. This line of questioning needs to be applied to each data set being deployed.

Expert systems: 'The digital tutor'

Introduced around 1965 by Feigenbaum as part of the Stanford Heuristic Programming Project, an expert system is a software program that emulates human decision-making processes by using databases to make expert decisions. Used extensively in online tutoring systems (as outlined in Yang and Zhang, 2019), it represents a system of diagnosing and remediating shortcomings in a student's or a teacher's knowledge base. However, this would always be limited by the ability of an individual system to update its own knowledge, as well as to transcend cultures (Mohammed and Watson, 2019). Examples in common use in English-speaking schools are the Education Perfect platform, Mathletics and Spellodrome, which offer personalized learning solutions to schools, complementing classroom activities. Schools that can afford access to these platforms are able to accelerate their children's learning, but those without access to the same resources may see their students fall behind, in relative terms.

Social robotics: 'The digital classroom assistant'

Robotics is an interdisciplinary branch of engineering that develops physical apparatus that can be used as a substitute for humans. In popular culture, we frequently see robots as either a revolutionary technical fix or representing a threat, perhaps by replacing teachers. This has been reflected in social attitude surveys, for example a recent EU investigation into public attitudes towards robots (European Commission, 2012). In reality, we see slow progress towards what are called 'social robots', via the use of neuromorphic computing techniques (see Mead, 1990, for the genesis and detailed description of the term) emulating the human brain in a similar manner to neural networking. The likelihood is that social robots are unlikely to be of use other than in limited knowledge and/or task domains, not least of all because of public resistance (Belpaeme *et al.*, 2018). There is also potential for 'cloud robotics', otherwise known as networked robots, a process in which educational robots share roles and policies via remote servers (Kehoe *et al.*, 2015), which could be useful in the context of school administration or educational assessment, for example. In turn, this links to the idea of an internet of things, whereby physical objects interconnect via embedded computing devices. This is leading to a growth in personalized learning, attendance trackers, school environmental controls and so on, although currently personalization in particular tends to be technology- and business-driven (Kucirkova, 2017), rather like the use of biometrics described above.

The first part of this article categorized artificially intelligent systems, also laying out a range of intended and unintended consequences for education. These have been examined both at the level of individual learners and institutions (such as through enabling new forms of curricular engagement, in the case of deep learning) as well as in a broader societal context (such as by triaging school admissions processes through predictive analysis). The next part of the article draws on the theories of both Bernstein (2000) and Homans (1958) to examine the social implications of AI for education in more depth.

Balancing competing social imperatives

As developments in AIED gain momentum, it is increasingly resembling a vast and somewhat ungainly landscape, as different technological and commercial approaches compete for attention. Relationships between learners and their schools are being redefined and reimagined, in order to render users governable through the medium of remote commercial systems, often based on subscription models. Any changes are generally being presented as a route to modernity. Indeed, the quotation from the Pearson (2018, emphasis added) press release that introduced this article is a typical example of the framing of AIED as uniquely transformational and forward-thinking: 'there are great opportunities to *improve learning outcomes* and to *enable better teaching* ... Pearson is committed to *transforming* the learning experience.' As Selwyn (2011b) makes clear, this is a rhetorical device that mirrors previous developments in educational technology.

This type of framing is based on the idea that education is perpetually in a receptive mode, almost exclusively impacted by external forces (in this case, commercial research and development). However, Biesta (2013) and others have argued that education itself also plays a key role in shaping the nature of the society in which we live. In this way, it can set the tone for the role of the individual within a broader societal context. This means that the relationship between AI as a technical field and education

itself has the potential to be significantly more dyadic than top-down, with influence continually travelling backwards and forwards. In order for the dyadic process to be truly successful, however, it requires a fine balance between the rights of the individual and the expectations of wider society. The theoretical work of Bernstein (2000) and Homans (1958) is very useful in demonstrating how this might be done.

Bernstein and pedagogic rights

Looking first at the work of Bernstein (2000), we can make use of his heuristic for pedagogic rights, which provides a useful counterpoint for challenging the indiscriminate use of ‘future needs’ as a blanket justification for AIED policies. These pedagogic rights have three aspects – enhancement, participation and inclusion – each of which can be related to developments in AIED. (I do not deal with Bernstein’s related concepts of ‘conditions’ or ‘levels’ here on grounds of length, and also because pedagogic rights seem to be the most appropriate in the case of considering AIED and fairness.)

Enhancement

Bernstein (2000: xx) conceptualized a right to the means of critical enhancement, described as ‘the means of critical understanding and to seeing new democracies’, implying a transformational process in which a learner can achieve a more open mind.

In terms of AIED, this is typically invoked rather vaguely, in the sense of ‘improving learning outcomes’, but also mediated by subscription-based models resulting in differential access to resources. This may limit the range of possible futures for some students, and they may not even be aware that this is happening. Any lack of transparency in the algorithmic profiling and monitoring of students therefore means that a greater degree of explanation and collaboration is required if there really is to be maximum enhancement for all. As is evident in the case of the subscription-based online textbook discussed in the ‘Legislating for fairness’ section, to take one example, it is hard to see how an equal degree of open-mindedness can be achieved if it is contingent on whether schools and parents can afford to pay for enhanced services in some cases and not in others.

Participation

Participation refers to social, cultural and intellectual involvement at a personal level. In other words, it means the right to engage in civic practice. Bernstein (2000: xx) describes this as ‘the right to be separate, autonomous’, as opposed to an individual being subsumed within a system, as is made clear by Frandji and Vitale (2016). Individual perspectives need to thrive, and there has to be scope for structures and systems to be sufficiently challenged by those who are subject to them.

In relation to AIED, this indicates that developments should fully involve users at all of these levels, ideally allowing them a fully collaborative role in the creation of systems. This allows ownership to be spread socially, culturally and intellectually across society, strengthening the basis for their operation. This is in contrast to systems being transient commercial artefacts to which particular groups subscribe (or to which they are required to subscribe). Such systems may become introspective and self-serving over time, as they only learn via a relatively limited population, which might be geographically or socially bound. An example is the forms of educational surveillance embedded in neural network systems, discussed earlier, where systems are trained in identifying and tracking students. Without collaborative involvement in

the development and use of such systems, outcomes are at risk of compromise by identifying false positives and negatives and compounding the deleterious impact of any privacy violations.

Inclusion

Bernstein (2000) defined 'inclusion' as a civic right to be involved in processes whereby order was constructed, maintained and changed. This represents a means of challenging the monopolistic/monolithic dominance of commercial (or even governmental) providers. Within this category, it is possible to envisage representative democratic structures growing out of the broadest possible user base.

This is likely to make AIED systems increasingly relevant over time rather than relying on the commercial judgements of remote organizations to decide what might be best for any particular group. It particularly applies in the case of minority populations. The discussions earlier in the article on machine learning and predictive analysis are of most interest in this regard. Bernstein's (ibid.) concept of inclusion means everyone should potentially have a seat at the table in terms of deciding what and how AIED systems are used within their community.

Each of the three pedagogic rights offers a unique perspective on the relationship between citizen and producer with regard to AIED. Together they recognize the need for commonalities within learning provision, as well as seeking to address the individual needs of learners from different backgrounds. Framed this way, the introduction of AIED is presented as a democratic project grounded in the idea of fairness, rather than purely a commercial trend.

The next section will consider the idea of fairness through the lens of Homans's (1958) social exchange theory, to explore the concept of altruism in AIED models.

Homans and social exchange theory

The work of Homans (1958) represents one of the first attempts at developing a social exchange theory (Blau, 1968; Ekeh, 1974; Leaton Gray, 2018). It links helpfully to the work of Bernstein (2000) for the purposes of this analysis. While many decades have passed since Homans's (1958) rudimentary model was developed, it still has relevance in the context of technological developments in education, and specifically in relation to the adoption of AI products and systems. This is because it differentiates between exchanges that are democratically equivalent, and those that are potentially imbalanced.

As described in relation to the adoption of commercial biometrics systems in schools (Leaton Gray, 2018), the introduction of any kind of profit motive shifts the fundamental nature of a social exchange. As Homans (1958) would argue, this results in one party seeking to achieve extraordinary benefits from the other, or, using the terms of the Pearson (2018) press release, to become the 'digital winner', implying a zero-sum game, with the loss of competitors equating to an exclusive advantage for the winning party. This is in contrast to a more altruistic positioning in which equal benefits are sought for both sides. Clearly, in marketing materials and press announcements, altruism is invoked as a justification for being a 'digital winner', and this is touched upon in the Pearson (ibid.) quotation as well. However, altruism and competition are not equivalent, and the terms 'authentic' and 'inauthentic' transactions can be usefully deployed here in order to differentiate between them (in a democratic sense rather than a technological one).

Table 2: Overt and covert forms of power among educational stakeholders

Aspect	Examples	Form of power	Data-related examples
Global	Markets and commercial activity; international educational organizations	Covert/overt	Big data analytics; socio-economic classification; data financialization
National	Government policy and curricula; inspection and regulatory bodies	Overt	Compliance and policy implementation
Community	Schools and colleges; teachers and students	Overt	Learning personalization; student monitoring and tracking
Family	Parents and household	Covert/overt	Collection of usage data and social surveillance

An *authentic* transaction in this context represents something that has equal engagement from both sides. This might mean adopting an artificially intelligent system that has been sought out by students, teachers and parents, and one in which they all have a role in the development process, as well as in sharing all the eventual benefits, whether these are financial or societal. This is in contrast to an *inauthentic* transaction, in which there is an imbalanced relationship. An example of this might be charging schools, parents or local education authorities differential fees for enhanced AIED resources, even though they have already effectively given away valuable data for free under the ‘legitimate interest’ category of privacy law, which provided the basis for the resource to be developed in the first place.

This tension between authentic and inauthentic transactions is best understood through considering wider power relationships. As identified in Table 2, stakeholders in education enjoy both overt and covert forms of power in relation to engaging with AIED systems, at different stages in their use.

It is when addressing differential forms of power that regulation can assist in codifying and reinforcing the rights of the citizen as consumer, and the role of the provider. New forms of privacy legislation have been designed to address this tendency towards knowledge commercialization and monopolization, through highlighting issues of consent and transparency in particular. This offers some insight into the key areas of concern that might usefully be considered. They are discussed in the next section.

Legislating for fairness: Privacy issues and GDPR

Central to the issue of perceived fairness of AIED systems are data privacy issues. Although school-related data use has been described as ‘overt’ in Table 2, in the sense that it is not secretive or confidential, this does not necessarily mean that students and their parents are fully cognizant of the extent to which children’s personal data are being used at school. Even though these data are often aggregated, for example to produce progress reports for particular classes, it is still possible to trace the attainment and engagement of individual students within that. In some ways, this does not matter, as parents and students no doubt expect schools to be keeping an eye on student progress and attainment in different ways. However, as AIED development accelerates, led by commercial organizations, there are more significant data privacy issues emerging. Analysing their relationship to the 2018 General Data Protection Regulations (GDPR) is a useful proxy for tracking the role of privacy generally in the development of AIED resources and platforms. Here, the law acts as a mirror to current societal concerns.

For example, a new development in the field of publishing is the digital textbook, hired from the publisher on a short-term basis for a year or two, using a subscription model. This allows a student to engage with the textbook material in a time-limited sense, tied to the educational institution he or she attends, which has paid for the subscription on the student's behalf, with built-in obsolescence. (See Blume and Ceasar, 2013, for an example of the serious impact of this obsolescence on school budgets in Los Angeles.) Behind the scenes, while the student is engaging with the textbook, the publisher is in turn collecting data on the way students are interacting with the material – for how long things are read, how attentively they are read, whether material needs to be repeated and so on. Yet students are unlikely to appreciate the fact that globalized publishing companies are collecting these data. The practice is something that is unlikely to benefit individual students in the short term, but it may allow for increased productivity by others in the future. In turn, this has the potential to lead to commensurately increased profits for the publisher as materials and algorithms are improved on the back end of this, leading to sophisticated AI systems that offer high levels of personalized learning via digital tutoring systems. On the surface, this seems like a positive development. However, over time, without sufficient controls in place, this could increase the kind of monopolistic pressures in the market that start to favour individual companies, rather than developing materials that might really be in the best interests of all learners.

Among other things, data protection law is designed to mitigate opaque uses of personal data when there is significant asymmetry between the data subject and a commercial company. In other words, it is there to promote fairness. Yet it is not sector specific, which means AIED presents a unique problem in this regard. Particular privacy issues need to be addressed if AIED is to represent a democratization of learning rather than purely a mercantile model of supply and demand. These issues are:

- the nature of consent
- transparency of processing
- the relationship between provider and student.

The nature of consent

There is a frequent misapprehension that consent is the primary mechanism for providing legal personal data processing under GDPR (European Union, 2016: Articles 6 (1) and 6 (2) and Recital 40). In reality this is not true – there are six lawful grounds for processing personal data, and consent is only one of these. The others are contract, legal obligation, vital interests, public task and legitimate interests (ICO, n.d.a). In the example of the digital textbook, the publisher would be able to argue that there is a legitimate interest in allowing for academic analytics to take place behind the scenes in order to develop AIED models. Individual consent by students may not be required, and the publisher may be allowed to dominate.

Transparency of processing

One of the new GDPR requirements is transparency of processing (European Union, 2016: Articles 13 and 14). The explanations of what personal data are being used, for what purpose and so on have to be expressed in a concise, plain and simple manner for data subjects. This is where AI systems potentially breach expectations of students, teachers and parents. Article 13 demands that suppliers describe any information

that has to be provided. Yet, under GDPR, there is also the need to tell people if automated decision making is being used, including profiling. Suppliers also need to provide meaningful information about the logic involved, as well as the significance and potential consequences of this kind of processing for the data subject. Yet how can this be achieved in the case of digital textbooks, for example, when revealing the way in which the algorithm works means revealing commercially sensitive information? It is hard to know how much detail can be reasonably expected. Currently there is no case law or significant supervisory authority guidance on how much detail will be required in this area, although what has been made clear is that the definition of profiling is going to be extremely wide, and that this will probably be included (ICO, n.d.b). We should therefore also acknowledge that there may be a clash between intellectual property law and the reasonable requirements of a company that has invested in algorithm development, against an absolute form of transparency in how they achieve their goals. There is an obvious tension here between the requirement for transparency and the essential workings of AIED.

Relationship between provider and student

From the examples throughout this article, we have seen that a significant consequence of AIED is that it involves an essential change in social relations between provider and student, and that this manifests itself through approaches to the scaling up of data collection and analysis, as well as its increasingly remote handling. As far as GDPR is concerned, if the issue of consent is not central to the acquisition and use of data by large international organizations, we need to consider the issues of public task/legitimate interest as grounds for the data processing taking place. This is not the same as schools processing attainment statistics themselves and releasing them to the government as a means of establishing how they are performing over time, for example. In such cases, the rights of governors, teachers and education authorities to this information is usually seen as overriding the rights of the individual. In the case of products developed by supranational companies and organizations, the conduct of this public task/legitimate interest balancing test instead moves away from local schools and teachers, and towards commercial or quasi-commercial organizations. Teachers and elected representatives are no longer able to decide if and how individual students might be tracked. Instead, tracking decisions are taken by parties unknown to children and parents. In this way, monitoring decisions move out of the direct hands of the teacher, to be controlled by third parties.

Knowledge asymmetry and fairness

On one level, the quotation from the Pearson (2018) press release at the beginning of the article simply presents us with a fairly obvious example of the natural tendency towards monopoly situations within a commercial environment. After all, the company with the largest database when the proverbial music stops will enjoy the most ability to influence educational developments to the organization's long-term advantage (although it may consider itself to have altruistic motives in addition to commercial ones, as Pearson (*ibid.*) makes clear).

However, there are distinct implications for democracy as this process unfolds. This is because in developing artificially intelligent systems, the concept of an educational population is moving beyond individual institutions, regions and even nation states, and becoming more globalized. As part of this process, stakeholders are proliferating, and data are being aggregated flexibly and at considerable scale

(Parks, 2014). One potential casualty here is policy accountability, as governments cede data autonomy to multinational commercial organizations (Grek, 2010). In this way, education policy starts to become increasingly removed from the individuals involved in the learning process, whose ability to influence the direction of development in a fully democratic sense, enabling their pedagogic rights, is correspondingly reduced.

This process is potentially compounded by differences in the epistemological basis for any data sets that provide the foundation for artificially intelligent systems. There is a danger that this is derived from the technocratic assumptions and expectations of the supranational model within which many stakeholders are working (Tan and Dimmock, 2019; Williamson, 2016). This imposes overt and covert forms of power, depending on different situations (see Table 2), something that may not be immediately obvious to a citizen attempting to ask the question, 'How fair are these systems?'. For example, within a proprietary expert tutor system, are you being shown a particular learning pathway because it is really the most suitable one for you, or is it because your local education authority did not have sufficient funds for a more varied one this academic year? Are you being allocated to a particular school purely on grounds of travel time, or is it because a predictive analysis system has been set to include or exclude particular housing developments nearby? Has a lower bar been set for sending a disciplinary letter home to your parents than for members of some other groups of students, because a machine learning system you do not know about identified you in a particular way? What is the precise nature of any data being shared about you? Without sufficient transparency, there is a danger that AIED algorithms subject users to a particular model of the ideal learner, or *ideal type*, to employ a sociological term, rather than fully accommodating fringe cases. This is because social distance from an ideal type can mean that fringe cases (for example, a certain type of student in a particular region) are increasingly overlooked as the quest for results across ever-larger populations becomes dominant. This can lead to inadvertent discrimination in multiple domains for particular individuals (Koene, 2017; Jackson, 2018). Therefore, issues such as fairness, trust, reliability and validity have never mattered more.

Conclusion

Respecting the pedagogic rights of the individual is key to the future success of AIED. More specifically, it is in the grey area between authentic and inauthentic transactions that the future of AIED needs to be mapped out and regulated. When there is an exchange of data, this needs to be achieved within the context of a relationship that is mutually beneficial. Unless this is done, digital differentiation will compromise the integrity of AIED from its birth, blighting it via new forms of social discrimination while compounding others. The pedagogic rights framework provided by Bernstein (2000) offers a good starting point for further strengthening provision. By promoting the involvement of users in all ways and at all levels of AIED systems, it is possible to exploit the different technologies under consideration to their fullest advantage. It is clear that AIED needs to be a truly collective project rather than a largely commercially driven one, as at present, or one imposed on school communities without sufficient scrutiny, transparency or consent. Governments need to take a lead here, going beyond mere rhetorical flourishes invoking modernity and progress, and instead moving into a solid regulatory position in which social inclusion can be fully guaranteed and supported. If this is done, then we may reach extraordinary heights of human flourishing. If it is not done, then we can soon expect to see increasing societal fragmentation instead. We should choose our path wisely.

Acknowledgements

I would like to acknowledge the extremely useful insights provided by Rebecca Eynon, Ansgar Koene and Iram Siraj, as well as two anonymous reviewers, during the development of this article.

Notes on the contributor

Sandra Leaton Gray is Associate Professor at UCL Institute of Education. She is a former teacher and sociology of education specialist, with a special interest in social and ethical issues surrounding contemporary identity, biometrics, AI and algorithms. Sandra is a member of the Privacy Expert Group of the Biometrics Institute, a senior member of Wolfson College, University of Cambridge, and a member of the advisory board of defenddigitalme, a children's privacy rights organization.

References

- Abbott, D. (2014) *Applied Predictive Analytics: Principles and techniques for the professional data analyst*. Indianapolis, IN: Wiley.
- Algorithm Watch (n.d.) 'Education, stock trading, cities and traffic'. Online. <https://tinyurl.com/ycjk6buu> (accessed 5 May 2020).
- Anderson, C. (2008) 'The end of theory: The data deluge makes the scientific method obsolete'. *Wired*, 23 June. Online. www.wired.com/2008/06/pb-theory/ (accessed 17 July 2019).
- BBC News (2017) 'Artificial intelligence school inspections face resistance'. BBC News, 21 December. Online. www.bbc.co.uk/news/technology-42425959 (accessed 17 July 2019).
- Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B. and Tanaka, F. (2018) 'Social robots for education: A review'. *Science Robotics*, 3 (21), Article eaat5954, 1–9. Online. <https://tinyurl.com/y88dorc4> (accessed 2 May 2020).
- Bernstein, B. (2000) *Pedagogy, Symbolic Control and Identity: Theory, research, critique*. Rev. ed. Lanham, MD: Rowman and Littlefield.
- Biesta, G. (2013) 'Responsive or responsible? Democratic education for the global networked society'. *Policy Futures in Education*, 11 (6), 733–44.
- Blau, P.M. (1968) 'The hierarchy of authority in organizations'. *American Journal of Sociology*, 73 (4), 453–67.
- Blume, H. and Ceasar, S. (2013) 'iPad software licenses expire in three years, LA Unified says'. *Los Angeles Times*, 19 November. Online. <https://tinyurl.com/y7crtl2u> (accessed 2 May 2020).
- boyd, d. and Crawford, K. (2012) 'Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon'. *Information, Communication and Society*, 15 (5), 662–79.
- Buolamwini, J. and Gebru, T. (2018) 'Gender shades: Intersectional accuracy disparities in commercial gender classification'. *Proceedings of Machine Learning Research*, 81, 1–15.
- Chorzempa, M., Triolo, P. and Sacks, S. (2018) *China's Social Credit System: A mark of progress or a threat to privacy?* (Policy Brief 18-14). Washington, DC: Peterson Institute for International Economics. Online. www.piie.com/system/files/documents/pb18-14.pdf (accessed 29 July 2019).
- Clarke, J., Newman, J., Smith, J., Vidler, E. and Westmarland, L. (2007) *Creating Citizen-Consumers: Changing publics and changing public services*. London: SAGE Publications.
- Cossins, D. (2018) 'Discriminating algorithms: 5 times AI showed prejudice'. *New Scientist*, 12 April. Online. <https://tinyurl.com/ybxbqtxj> (accessed 4 May 2020).
- Ekeh, P.P. (1974) *Social Exchange Theory: The two traditions*. Cambridge, MA: Harvard University Press.
- Elgammal, A. (2019) 'AI is blurring the definition of artist: Advanced algorithms are using machine learning to create art autonomously'. *American Scientist*, 107 (1), 18–21.
- European Commission (2012) *Public Attitudes towards Robots* (Special Eurobarometer 382). Brussels: European Commission.
- European Union (2016) 'Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General

- Data Protection Regulation) (Text with EEA relevance). Online. <http://data.europa.eu/eli/reg/2016/679/2016-05-04> (accessed 3 September 2019).
- Eynon, R. (2013) 'Editorial: The rise of big data: What does it mean for education, technology, and media research?'. *Learning, Media and Technology*, 38 (3), 237–40.
- Feenberg, A. (1999) *Questioning Technology*. London: Routledge.
- Frandji, D. and Vitale, P. (2016) 'The enigma of Bernstein's "pedagogic rights"'. In Vitale, P. and Exley, B. (eds) *Pedagogic Rights and Democratic Education: Bernsteinian explorations of curriculum, pedagogy and assessment*. London: Routledge, 13–32.
- Goldhaber-Fiebert, J.D. and Prince, L. (2019) *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County's Child Welfare Office*. Online. <https://tinyurl.com/y9vmk8kz> (accessed 12 May 2020).
- Good, I.J. (1966) 'How to estimate probabilities'. *Journal of the Institute of Mathematics and Its Applications*, 2 (4), 364–83.
- Grek, S. (2010) 'International organisations and the shared construction of policy "problems": Problematisation and change in education governance in Europe'. *European Education Research Journal*, 9 (3), 396–406.
- Har Carmel, Y. (2016) 'Regulating "big data education" in Europe: Lessons learned from the US'. *Internet Policy Review*, 5 (1), 1–17. Online. <https://tinyurl.com/ycmjrqe8> (accessed 2 May 2020).
- Herodotou, C., Hlosta, M., Borooowa, A., Rienties, B., Zdrahal, Z. and Mangafa, C. (2019) 'Empowering online teachers through predictive learning analytics'. *British Journal of Educational Technology*, 50 (6), 3064–79.
- Hess, H. (2019) 'Privacy vs modernity: Artificial intelligence in China's infrastructure'. In Blowers, M., Hall, R.D. and Dasari, V.R. (eds) *Disruptive Technologies in Information Sciences II* (SPIE Proceedings 11013). Online. <https://tinyurl.com/yb5m9494> (accessed 5 May 2020).
- Homans, G.C. (1958) 'Social behavior as exchange'. *American Journal of Sociology*, 63 (6), 597–606.
- Hornigold, T. (2018) 'The first novel written by AI is here – and it's as weird as you'd expect it to be'. *SingularityHub*, 25 October. Online. <https://tinyurl.com/y88xpv13> (accessed 2 May 2020).
- Huang, A. and Wu, R. (2016) 'Deep learning for music'. Online. <https://arxiv.org/pdf/1606.04930.pdf> (accessed 5 May 2020).
- ICO (Information Commissioner's Office) (n.d.a) 'Rights related to automated decision making including profiling'. Online. <https://tinyurl.com/ydxccc5a> (accessed 12 May 2020).
- ICO (Information Commissioner's Office) (n.d.b) 'Lawful basis for processing'. Online. <https://tinyurl.com/y24ybmqf> (accessed 12 May 2020).
- Jackson, J.R. (2018) 'Algorithmic bias'. *Journal of Leadership, Accountability and Ethics*, 15 (4), 55–65.
- Jarke, J. and Breiter, A. (2019) 'Editorial: The datafication of education'. *Learning, Media and Technology*, 44 (1), 1–6.
- Kehoe, B., Patil, S., Abbeel, P. and Goldberg, K. (2015) 'A survey of research on cloud robotics and automation'. *IEEE Transactions on Automation Science and Engineering*, 12 (2), 398–409.
- Kirkman, J. (2014) 'Building a culture of trust: Trust in the use of educational technology'. *Australian Educational Computing*, 29 (1), 1–10. Online. <https://tinyurl.com/y8cqrzcp> (accessed 2 May 2020).
- Koene, A. (2017) 'Algorithmic bias: Addressing growing concerns'. *IEEE Technology and Society Magazine*, 36 (2), 31–2.
- Kucirkova, N. (2017) *Digital Personalization in Early Childhood: Impact on childhood*. London: Bloomsbury Academic.
- Lawn, M. (2013) 'Introduction: The rise of data in education'. In Lawn, M. (ed.) *The Rise of Data in Education Systems: Collection, visualisation and uses*. Oxford: Symposium Books, 7–10.
- Leaton Gray, S. (2018) 'Biometrics in schools'. In Deakin, J., Taylor, E. and Kupchik, A. (eds) *The Palgrave International Handbook of School Discipline, Surveillance, and Social Control*. Cham: Palgrave Macmillan, 405–24.
- Lehr, P. (2019) 'Undemocratic means: The rise of the surveillance state'. In Lehr, P. *Counter-Terrorism Technologies: A critical assessment*. Cham: Springer, 169–79.
- Luckin, R. (2018) *Machine Learning and Human Intelligence: The future of education for the 21st century*. London: UCL Institute of Education Press.
- Mackenzie, A. (2017) *Machine Learners: Archaeology of a data practice*. Cambridge, MA: MIT Press.
- Manolev, J., Sullivan, A. and Slee, R. (2019) 'The datafication of discipline: ClassDojo, surveillance and a performative classroom culture'. *Learning, Media and Technology*, 44 (1), 36–51.
- McCarthy, J., Minsky, M.L., Rochester, N. and Shannon, C.E. (1955) *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Hanover, NH: Dartmouth College.
- Mead, C. (1990) 'Neuromorphic electronic systems'. *Proceedings of the IEEE*, 78 (10), 1629–36.

- Mohammed, P.S. and Watson, E. (2019) 'Towards inclusive education in the age of artificial intelligence: Perspectives, challenges, and opportunities'. In Knox, J., Wang, Y. and Gallagher, M. (eds) *Artificial Intelligence and Inclusive Education: Speculative futures and emerging practices*. Singapore: Springer, 17–37.
- Ofsted (2018) *Methodology Note: The risk assessment process for good and outstanding maintained schools and academies*. London: Ofsted. Online. <https://tinyurl.com/yaqmt8yx> (accessed 12 May 2020).
- Ong, V.K. (2016) 'Business intelligence and big data analytics for higher education: Cases from UK higher education institutions'. *Information Engineering Express*, 2 (1), 65–75.
- Pardo, A. and Siemens, G. (2014) 'Ethical and privacy principles for learning analytics'. *British Journal of Educational Technology*, 45 (3), 438–50.
- Parks, M.R. (2014) 'Big data in communication research: Its contents and discontents'. *Journal of Communication*, 64 (2), 355–60.
- Pearson (2018) 'Pearson hires new head of artificial intelligence'. Press release, 2 July. Online. <https://tinyurl.com/y7nhlz64> (accessed 12 May 2020).
- Polonetsky, J. and Tene, O. (2014) 'The ethics of student privacy: Building trust for ed tech'. *International Review of Information Ethics*, 21, 25–34.
- Reynolds, M. (2017) 'UK's Nudge Unit tests machine learning to rate schools and GPs'. *Wired*, 14 December. Online. <https://tinyurl.com/ybjjka98> (accessed 3 May 2020).
- Russell, S. and Norvig, P. (2016) *Artificial Intelligence: A modern approach*. 3rd ed. Harlow: Pearson Education.
- Samuel, A.L. (1959) 'Some studies in machine learning using the game of checkers'. *IBM Journal of Research and Development*, 3 (3), 210–29.
- Slater, N. and Bailey, P. (2018) 'Code of practice for learning analytics'. Online. <https://tinyurl.com/ya6cz23d> (accessed 3 May 2020).
- Selwyn, N. (2011a) "'It's all about standardisation": Exploring the digital (re) configuration of school management and administration'. *Cambridge Journal of Education*, 41 (4), 473–88.
- Selwyn, N. (2011b) *Education and Technology: Key issues and debates*. London: Continuum.
- Selwyn, N. (2015) 'Data entry: Towards the critical study of digital data and education'. *Learning, Media and Technology*, 40 (1), 64–82.
- Takahashi, I., Oki, M., Bourreau, B., Kitahara, I. and Suzuki, K. (2018) 'FUTUREGYM: A gymnasium with interactive floor projection for children with special needs'. *International Journal of Child-Computer Interaction*, 15, 37–47.
- Tan, C.Y. and Dimmock, C. (2019) 'National and transnational influences on school organization'. In Connolly, M., Eddy-Spicer, D.H., James, C. and Kruse, S.D. (eds) *The SAGE Handbook of School Organization*. London: SAGE Publications, 414–29.
- Tudor, J. (2015) 'Legal implications of using digital technology in public schools: Effects on privacy'. *Journal of Law and Education*, 44 (3), 287–343.
- Turhan, M., Erol, Y.C. and Ekici, S. (2016) 'Predicting students' school engagement using artificial neural networks'. *International Journal of Advances in Science, Engineering and Technology*, 4 (2), 159–62.
- Turing, A.M. (1950) 'Computing machinery and intelligence'. *Mind*, 59 (236), 433–60.
- 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.
- Williamson, B. (2016) 'Digital education governance: Data visualization, predictive analytics, and "real-time" policy instruments'. *Journal of Education Policy*, 31 (2), 123–41.
- Williamson, B. (2017a) 'Who owns educational theory? Big data, algorithms and the expert power of education data science'. *E-Learning and Digital Media*, 14 (3), 105–22.
- Williamson, B. (2017b) 'Decoding ClassDojo: Psycho-policy, social-emotional learning and persuasive educational technologies'. *Learning, Media and Technology*, 42 (4), 440–53.
- Yang, J. and Zhang, B. (2019) 'Artificial intelligence in intelligent tutoring robots: A systematic review and design guidelines'. Online. <https://tinyurl.com/y7kqgojg> (accessed 17 July 2019).