

## Can we avoid digital structural violence in future learning systems?

Niall Winters, Rebecca Eynon, Anne Geniets, James Robson and Ken Kahn  
Learning and New Technologies Research Group  
Department of Education, University of Oxford

### Abstract

This paper introduces the concept of digital structural violence and examines the negative role it could have in future learning systems. To address it, we propose a new interdisciplinary research agenda at the intersection of three current but disparate lines of work that:

1. Use the concept of epistemic privilege to theorise the inclusion of marginalised learners in the design of learning systems, and utilise participatory action research and emancipatory methodologies to pragmatically ensure this happens;
2. Support young learners and teachers to understand and build their own artificial intelligence algorithms;
3. Develop sustainable interdisciplinary links with computer science to address digital structural violence at the algorithmic level and to make its societal implications and underlying processes more widely understood, especially by teachers.

Taken together, these provide for a material form of resistance to digital structural violence and a theoretically and methodologically coherent future research agenda for building just learning systems.

### 1. Introduction

The aim of this paper is to discuss the potential negative consequences of artificial intelligence in learning, through a novel conceptualisation we term *digital structural violence*. We propose a research agenda on artificial intelligence (AI) in future learning systems. This is designed specifically to address the needs of the most marginalised, as they are the group who will suffer the most from digital structural violence. By the 2020s, many are predicting that AI will be commonplace in technology enhanced learning (Luckin et al., 2016) and although the future is likely to be complex, we believe that AI will likely be a very relevant consideration in the learning landscape. While predicting the future is always a fraught exercise, our examination of what digital structural violence might look like over the next 10 years is based on a concrete extrapolation from existing research on (a) digital inequality, (b) social exclusion from educational opportunities and (c) technical developments in artificial intelligence. While the precise future form of digital structural violence remains unknown, we hypothesise that it will be shaped by: (i) new AI algorithms; (ii) data that reflects and promotes the aims of those with power; (iii) decision-making algorithms that are increasing difficult to dissect but paradoxically will gain increased currency; and (iv) an ahistorical view that loses sight of AI's genesis as a force for good, particularly in the fields of children's programming (Kahn, 1977) and ubiquitous computing (Kidd et al., 1999)).

Currently, we are seeing a renaissance for the field of Artificial Intelligence in Education (AIEd) due to the considerable excitement around new machine learning techniques, which attempt to personalise learning in fundamentally new ways. To understand how common this approach has become, we can already see its use (and improvements over time) in many real

world applications today: for example, AI conversational agents have become common in phones and PCs (e.g. Siri, Cortana, Google Assistant); recommender systems are now a well-established feature on shopping websites; algorithms are being used to diagnose disease in some areas with similar accuracy to humans, digital surveillance is being used in schools and in public spaces, and driverless vehicles are on the horizon. How has this been made possible? A central reason is that the availability of vast amounts of data (e.g. hundreds of millions of photos are uploaded to Facebook every day) has enabled the development of powerful new AI algorithms can make predictions for an individual based on examining patterns in these very large datasets. Essentially neural networks are used to train a ‘machine intelligence’ with lots of examples from which it can learn (Vandewaetere, Desmet & Clarebout, 2011). This is sometimes called ‘deep learning’ and differs from the previously dominant rules-based approach in AIED, which relied upon the development of complex learner, domain and instructional models.

However, a critical disadvantage of pattern-based approaches is that they are open to serious biases as they reflect the data they were trained with. While there is an absence of studies of bias in research related to education and technology, this has been an active area of research in other sectors: For example, studies have highlighted gender and racial biases in systems for career development and systems built to predict future criminals (e.g. Angwin et al., 2016; Kay et al., 2015), with a 2017 paper in *Science* reporting that datasets often contain “accurate imprints of our historic biases, whether morally neutral as toward insects or flowers, [or] problematic as toward race or gender” (Caliskan, Bryson & Narayanan, 2017).

Facial recognition software has been shown to be significantly biased towards being able to recognise male faces, and men and women with the lightest skin tones (Buolamwini and Gebru, 2018); search engine results similarly privilege white men, and discriminate against people of colour, particularly women (Noble, 2018); and there are many concerns around the digital inequalities rendered around issues of gender (e.g. the default settings of Amazon’s female assistant Alexa) (Eynon, 2018).

Though the examples above have clear implications for informal learning, these questions have less often been asked within more formal Educational settings. From past work on the links between education, technology and inequality we know that inequalities can be reinforced or exacerbated by educational technologies (White and Selwyn, 2012; Funes and Mackness, 2018; Macgilchrist, 2019). While these may not be intentional, they often have a structural dimension, where advances in technologies benefit those who are already in a position of power. For example, in the case of young people, research has demonstrated that they vary significantly in their level of competence in using technology (Davies and Eynon, 2013; Hargittai, 2010). Those who are better off – and better connected – are able to make more meaningful use of technology, in contrast to those young people who experience constrained usage, leading to them sometimes being openly hostile to technology (Eubanks, 2012) due to the significant unequal social structures they face (Robinson, 2009; Eynon and Geniets, 2016).

To take another well-known example, Massive Open Online Courses (MOOCs) started out as a way to “democratise education” by providing free access to online university courses. However, in practice those who were supposedly going to benefit most from technology often ended up being least likely to do so. For example, research has detailed how “HarvardX registrants ... on average live in neighborhoods with median incomes approximately .45

standard deviations higher than the US population” (Hansen & Reich, 2015). Research shows that MOOCs now are a tool for universities to outsource online Masters programmes for professionals (Reich and Ruipérez-Viliente, 2019). Furthermore, there are concerns of forms of educational redlining, (i.e. the differential use of AI bots in MOOCs) with those able to pay provided with a more sophisticated human-mediated experience. Critically, although technologies are offered to all, how they are implemented and used differs depending on the interaction between an individual’s agency and social structure. Given what we know about widening digital inequalities, engendered by ill thought-out uses of learning technologies, it is tempting to predict that digital inequalities will continue to worsen as AI technology becomes more pervasive.

This is not a completely new problem, as the issue of digital inequality is seen again and again as new technologies gain prominence, reproducing disadvantage (Bourdieu, 1977). What is different this time around is that rather than existing social inequalities being reinforced by online activities (e.g. Otto et al., 2005; Drentea and Moren-Cross, 2005), it is the *extent to which* AI could bring multiple disadvantages together in a manner we have not seen before, across all life stages and across learning both formal and informal. It is the combination, for example, of biases in search engines, MOOCs and YouTube, alongside everyday injustice in health, education and employment that work together to compound disadvantage. This could be at a systematic and pervasive scale that is unavoidable. It is this we term *digital structural violence*.

## **2. Structural violence**

“Structural violence” is a term coined by Galtung (1969) to refer to the structural and institutional constraints placed on (marginalised) people, which prevent them from living the life they value. Any social structure, be it economic, political, educational or cultural, which prevents individuals from expressing themselves and developing their full potential, can accordingly be conceptualized as structural violence. Embedded in social structures, such violence is often normalised by stable institutions and reproduced within established cultural practices rooted in a range of political, legal and economic systems (DuNann Winter & Leighton, 2001). At its most extreme, structural violence leads to significant suffering and death through, for example, limitations placed on access to medicine or healthcare. However, violence can be embedded in a wide range of social structures that lead to varied forms of discrimination and disadvantage related to, for example, race, gender, age, socio-economic status (SES), geographical location etc. As argued by DuNann, Winter and Leighton, structural violence produces suffering just as much as direct violence, but the ‘damage is slower, more subtle, more common, and more difficult to repair’ (2001, p99).

Within an educational context, a great deal of work has been undertaken that highlights a range of ways in which embedded structures disadvantage different groups or individuals. For example, work in higher education consistently emphasises the ways in which educational system-wide structures disadvantage low SES and black, Asian and minority ethnic students’ access to university (O’Sullivan, Robson and Winters, 2018) and negatively shape both educational outcomes and labour market opportunities (Crawford et al., 2016). Educational structural violence can be conceptualised in pervasive terms across a wide range of subtle arenas. For example, increased marketisation across education systems and embedded school accountability measures can significantly shape the educational experiences and outcomes of low SES students or individuals with special educational needs, negatively shape the educational workforce (Robson, Randhawa, and Keep, 2018), and even lead to student

exclusion (Daniels and Cole, 2010). Structural violence embedded in institutional governance can lead to casualisation of certain sectors of the educational workforce leading to social, economic and mental health related consequences (Loveday, 2018).

As described above, technology in educational contexts frequently reproduces social structures and so has the capacity to reproduce the structural violence embedded within them and the cultural norms that legitimate that violence. However, technology also has the capacity to introduce new social, institutional, cultural, political, and economic structures with new forms of embedded structural violence, often at the expense of the most marginalised groups in society, exacerbating disadvantage. The increasing prominence of AI and its reliance on structuring algorithms has the potential to deepen these issues. It is therefore critical to examine existing mechanisms of digital structural violence in order to understand future developments.

## **2.1 Existing forms of structural violence: what we already know**

The immediate challenge to address is to ensure that learning technologies do not become another pervasive mechanism for structural violence. While, it is clear from what we have outlined above that learning systems can already occasionally be a mechanism of structural violence, we need to be very aware as to how we do not (inadvertently) intensify these mechanisms. Indeed, pressing discussions need to be had about how to build mechanisms of resistance that at the very least provide a minimum standard for the reconfiguration of learning systems as a social good.

We can start by building on what we already know from sociology and educational research over the past decade, which has placed increased scrutiny on the data and algorithms underlying AI approaches. A core aim of this work is to understand the social implications of technical processes, including the ways in which new forms of inequality may emerge from technical implementation. This can happen in two main ways that are relevant to this paper. The first is the privilege of experts, meaning that the ways in which algorithms are designed, implemented and tested is confined to small groups of (often homogenous) experts, who most likely share a common view of the world. Therefore, they may not be the best community to self-judge and correct their own biases or be entirely responsible for building technology for educational purposes. The second is the way in which data collected by technologies ends up better representing those who have access to, and frequently use, technology. In this way, existing power relations are reconstituted through these new technical processes. We now explain these two critical issues in more detail before going on to show how to address each of them.

### **2.1.1 The privilege of experts**

In response to concerns regarding the uptake in the use of Big Data, boyd and Crawford (2012) have discussed the need to question “the assumptions, values, and biases of this new wave of research”, p. 675). Drawing on Manovich (2011), they note the privileges of technical experts (often industry-based and colloquially now known as “tech bros”) who know how to analyse these vast datasets, allowing them “to determine the rules about how Big Data will be used, and who gets to participate” (p. 675). They caution researchers to be aware of the bias produced by this privilege, including the kinds of research that gets done. However, AI presents a new problem: the technical processes used to achieve an outcome (translation, for example) are implementable “without necessarily knowing why they work or being able to

show the logic behind a system's decision" (Ananny & Crawford, 2016, p. 9). Computer scientists know how to design AI systems but the "internal logic" of the algorithms never takes an observable state. This presents a serious challenge for developing critical consciousness to address exclusionary decisions that affect marginalised learners disproportionately. Fortunately, recent work by AI researchers is seeking to address this important issue and we discuss this further in Section 3.

### **2.1.2 Social power is reconstituted through technical processes**

As Wessels points out, not having an online profile can lead to having less power within a 'digitally organised society'. This is because such a profile is required to provide certain kinds of 'status markers' that enable authentication. If such markers are not present, then individuals become routinely excluded from a range of online services in both the commercial and public sector (Wessels, 2015: 2801). This exclusion could potentially be exacerbated in a data saturated society. AI systems are now trained based on large datasets of real world uses of technology (e.g. social media feeds). Yet, not everyone in society is equally represented in these data sets. Those who are (or have the choice to be) well connected, or particularly active or extrovert online will be better represented in these large data sets on which AI systems are trained; and this has implications for learning and education as it does for other areas of society (Eynon, 2013).

Continuing this theme, Selwyn (2015) points out data's perceived advantages "reflect an underlying belief that digital data render social processes and social relations more knowable and, it follows, more controllable" (p. 66). He too discusses his concerns regarding digital data and the reproduction of inequalities, drawing our attention to issue of "social power and control being reinforced, or perhaps reconstituted, through data-driven processes" (p. 71). Drawing on Manovich (2011), Selwyn highlights the inequalities between people who 'do things' with data and people who have things 'done to them' with data. He calls for more research on differential experiences of digital data in education. We address this point further in Section 3.2.

## **2.2 Digital structural violence engendered by AI**

Revisiting our conceptualisation, digital structural violence is the violence that could occur when AI brings multiple disadvantages together at a systematic and pervasive scale. Inequalities are then exacerbated by unknowing AI algorithms and our concern is that future learning systems will be designed to promote forms of, and opportunities for, learning that is increasingly tailored to those with existing advantages. The mechanisms by which this may happen are already evident on social media and on sites such as YouTube, where informal learning increasingly takes place<sup>1</sup> (Pew Research 2018). James Brindle provides an informative insight<sup>2</sup> into how these mechanisms work. He examined providers' content development processes that were designed to generate high exposure (and thus views) on YouTube, aligned with their recommendation algorithms. He showed how children's video content is being automatically generated, which has led to unsuitable content being made widely available. What differentiates this example from what has come before is that the infrastructural design of YouTube: here the underlying algorithms are tuned for maximising revenue streams. Critically, this led to priority being given, not to the development of quality content, but the

---

<sup>1</sup> [https://ipkl.gu.se/english/Research/research\\_projects/online-instructional-videos-and-learning-of-practical-skills](https://ipkl.gu.se/english/Research/research_projects/online-instructional-videos-and-learning-of-practical-skills)

<sup>2</sup> <https://medium.com/@jamesbrindle/something-is-wrong-on-the-internet-c39c471271d2>

automatic generation of content for profit. As Brindle notes: “It’s not about trolls, but about a kind of violence inherent in the combination of digital systems and capitalist incentives”. Here, the way the system is designed (in this case YouTube) and the way data is generated at scale means that outputs are generated that promote a type of digital structural violence that has exposed children (in this case) to forms of engagement that is outside of their cultural norms, over which they have very little control. This has been driven by algorithmic incentivisation that has a ‘ranking culture’, which is seen to thrive on controversy and, as shown by Rieder, Matamoros-Fernández & Coromina (2018), “systematically beats out mainstream actors in terms of exposure” (p. 50).

Similarly, in the context of professional learning, increased scrutiny is being placed on the use of social media to support teachers’ professional development (e.g. Bergviken et al., 2018). The affordances of online social spaces like Twitter and Facebook are often emphasised in professional circles as providing teachers with agency to take control of their own professional development by creating their own educational networks through personalised feeds and grassroots-based digital discussions (Miller, 2017). However, as highlighted by Robson (2016 and 2018), personalised feeds and social networks can be significantly shaped by dominant discourses and the political and economic agendas of key stakeholders within online social spaces. Most importantly, as algorithm-driven personalisation increasingly dominates social media, users’ personal feeds are all too often shaped by their browsing and particularly online shopping habits. Displayed content can therefore be a combination of posts that are similar to those a user has previously engaged with (reproducing dominant discursive structures) and posts driven by the advertising agendas of commercial partner. The undermining of professional agency through such embedded and algorithmic structures can be conceptualised as a form of structural violence that negatively shapes professional knowledge, practice, and identity.

Within more structured and formal educational contexts, Williamson has highlighted the ways in which digital data drive university governance in ways that include both human and non-human actors and increasingly algorithm-driven processes (2018). This work highlights the sociotechnical networks that underpin educational infrastructure that carry embedded political, economic and social agendas driven by forces of marketisation. Such digital structures shape both students’ and university workers’ lives and form part of a wider framework of technologies of power and digital surveillance (Zuboff, 2019; Eubanks, 2018; O’Neil, 2016) and, as such, can be conceptualised in terms of digital structural violence.

### **3. New approaches to tackling digital structural violence**

It is possible that as AI becomes increasingly normalised, digital structural violence will *not* be seen as deviant but instead will be “defined as moral in the service of conventional norms and material interests” (Bourgois and Scheper-Hughes, 2004, p. 318). It is therefore important that as researchers and practitioners, we confront the challenges of contributing to building an alternative future, where the disadvantages unleashed by digital structural violence are addressed as early as possible.

In order to examine how we build the material and conceptual resources to combat digital structural violence, we suggest three potential avenues of future research:

1. Use the concept of epistemic privilege to theorise the inclusion of marginalised learners in the design of learning systems. To pragmatically support this, we suggest using participatory action research and emancipatory methodologies to pragmatically ensure this happens;
2. Support young learners and teachers to understand and build their own artificial intelligence algorithms;
3. Develop sustainable interdisciplinary links to computer scientists to address digital structural violence at the algorithmic level and to make the societal implications and underlying processes more widely understood, especially by teachers.

We suggest that these avenues of research are considered in light of broader approaches to the socio-political landscape of technological innovation. These have a long history; for example, in 1972 Joseph Weizenbaum wrote an article in *Science*, in which he discussed the “impact of the computer on society” (Weizenbaum, 1972). He pointed out that:

“... the computing metaphor is as yet available only to an extremely small set of people. Its acquisition and internalization, hopefully as only one of the many ways to see the world, seems to require experience in program composition, a kind of computing literacy. Perhaps such literacy will become very widespread in the advanced societal sectors of the advanced countries. But, should it become the dominant mode of thinking and be restricted to certain social classes, it will prove not merely repressive in the ordinary sense, but an enormously divisive societal force. For then classes which do and do not have access to the metaphor will, in an important sense, lose their ability to communicate with one another.” (p. 613-614)

This point remains very relevant today, perhaps even more so, as we grapple with the wider negative implications of an AI future. For example, a 2018 special issue of the *Philosophical Transactions of the Royal Society A*, considered the ethical, legal, and technical opportunities and challenges in governing AI in an “accountable, fair and transparent” manner. It is within this emerging context that we put forward our three area of research focus. In doing so, we do not seek to undersell the scale of the task. The current neo-liberal logics at the heart of education systems across the world and wider structural inequalities can constrain significantly the actions of individuals. However, like other critical scholars we argue that the responsibility for social change lies with everyone: students, teachers, tech developers, researchers and policy makers - both in the use and design of technology. This change can be achieved through the actions of small groups and individuals. The proposals detailed in Sections 3.1-3.3 are a starting point for such actions.

Our arguments build on past and ongoing commitments to critical scholarship. Scholars outside of the field of Education have argued for moving beyond a simple critique of the current direction of travel in data and AI practices, towards a more action-based approach that suggests practical ways forward to provide more socially responsible outcomes (Daly et al. 2019). Supporters of this view recognise the challenges of local initiatives contributing to large scale social change, particularly when the current networks of power and privilege are entrenched. Nevertheless, they view such an approach as important for change to occur.

For example, Green argues for a framework for the education of data scientists to help them to “recognize themselves as political actors engaged in normative constructions of society and,

as befits political work, evaluate their work according to its downstream material impacts on people's lives" (Green, 2018:7). Green notes that academia does not necessarily value this kind of data work, policy and non-profits rarely engage in data work, and so data scientists often end up in the commercial sector with its neo-liberal value systems. Yet while the curriculum is only one part of a wider strategy for social change, it is an important one. Thus, Green details an approach to a curriculum for data scientists that could bring about positive change on an individual but also at the macro level: e.g. the use of placements, learning about STS, changing the demographics of who become data scientists, and changing the incentive structures of the field (e.g. changing what good research in this space looks like) (Green, 2018).

In a similar vein, D'Ignazio and Klein (2019) draw on feminist theory to argue for the importance of multiple actors needing to be involved in data work, that recognises and values the different views of those involved or implicated in its use, with particular effort and attention to those who are most marginalised. Such an approach can allow for a range of possibilities that may result in positive social change. For example, participatory mapping initiatives that embed community storytelling, facilitating data informed conversations within and across communities, and using data to support consensus building can all be used to actively tackle and make visible structural inequalities, highlighting how data is not objective and neutral, and is instead biased, benefiting those in power and harming others.

Like Green, D'Ignazio and Klein note that data projects may be easier in some domains (e.g. in activism, law and journalism) but may be more difficult in others because of structural issues such as a lack of funding for projects that actively take on and question existing power structures, and the dearth of incentives to those who have access to data to give this to others. This is because the current structural arrangements suit them, and that inequalities are often not apparent to those who are benefiting from them (D'Ignazio and Klein, 2019). Nevertheless, the authors believe these approaches can make a difference at a more structural level. As the authors note:

"A data scientist is not going to save democracy, but a well-designed, data-driven, participatory process that centers the standpoints of those most marginalized, empowers participants and builds new relationships across lines of social difference? Well, that might just have a chance." (D'Ignazio and Klein, 2019, n.p.)

We now move on to discuss our three avenues of future research in more detail.

### **3.1 Working with marginalised groups**

To being to address the future influence of digital structural violence, we need to start by working with learners who are most likely to be affected, which research shows are often the most marginalised (Farmer, 2004). In brief, we wish to focus on how to achieve this goal both theoretically and methodologically.

Feminist scholars in the early 1980s developed the concept of epistemic privilege (Harding & Hintikka, 1983; Hartsock, 1983; Narayan, 1988). It is the claim "that members of an oppressed group have a more immediate, subtle and critical knowledge about the nature of their oppression than people who are non-members of the oppressed group" (Narayan, 1988, p.35). Epistemic privilege has its roots in attempts to help people from different backgrounds in "working together across difference" (ibid, p.33). Narayan argues that we should not expect

the oppressed to have clearer knowledge of how they are oppressed as “oppression is often partly constituted by the oppressed being denied access to education and hence to the means of theory production, (which would include detailed knowledge of the history of their oppression, conceptual tools with which to analyze its mechanisms etc.) [and] the oppressed may not have a detailed causal/structural analysis of how their specific form of oppression originated, how it has been maintained and of all the systemic purposes it serves” (p. 36).

The implications of this are that researchers need to work together with marginalised learners to understand the effects and implications of digital structural violence. Methodologically, this can be achieved in a number of ways but we advocate here for the use of participatory methods (White et al, 2004) and emancipatory ethnographies (Pane and Rocco, 2009), while keeping in mind the power structures implicit in any such collaboration between researchers, practitioners and marginalised groups. The core aim is to privilege the language, perspectives, experience and capabilities of marginalised learners. This approach could also combine participatory action research with advanced data analytics and social network analysis as necessary, but as noted by Abdurahman the complexities of this are often overlooked: “it is not just that classification systems are inaccurate or biased, it is who has the power to classify, to determine the repercussions / policies associated thereof and their relation to historical and accumulated injustice?”<sup>3</sup>. Researchers need to work together with marginalised learners to co-develop a better understanding the biases that are already embedded in the structures of AI systems. One example is to look for biases in large datasets relevant to education (for example, from previously run MOOC or VLE modules) used for training AI algorithms and critically understand the power structures in place that enabled this to happen. Once identified, these can help researchers generate more representative datasets. Alongside this focus on data, is the need to explore the wider social, political and economic context that enabled MOOCs to become such a prevalent focus of discussions on widening participation, in order to also begin to develop implementation strategies for the use of AI in education that enable positive change. As work in this area develops, many complex issues will likely emerge. However, despite the likely challenges, these methodologies are best placed to tackle structural violence, and far more research of this nature is required.

### **3.2 Helping young people and teachers to understand and explore AI algorithms**

Every day there is news of new applications of machine learning that can enhance social life. Sometimes it is for entertainment as when SnapChat filters add animations or distort people's faces<sup>4</sup>. Sometimes it saves lives as when students used machine learning to predict wildfires<sup>5</sup>. These examples show the positive possibilities of AI. Addressing digital structural violence does not mean a rejection of AI, but rather recognition that AI is a complex social and cultural artefact that can be reconfigured and reshaped.

We believe that a key part of addressing digital structural inequalities is to expose young people and teachers to the positive as well as negative possibilities of AI as early as possible. The way we chose to develop and use AI is a choice that we as a society make. It has been long recognised that one fundamental way to achieve this is through Education. Indeed, AI for children already existed in 1977 (Kahn 1977). An important part of such an educational

---

<sup>3</sup> <https://medium.com/@blacksirenradio/fat-be-wilin-deb56bf92539>

<sup>4</sup> <https://www.youtube.com/watch?v=Pc2ajxnmzh0>

<sup>5</sup> <https://www.blog.google/technology/ai/fighting-fire-machine-learning-two-students-use-tensorflow-predict-wildfires/>

programme, is to enable young people to use AI to become better motivated and empowered to produce their own artefacts. Recently, interesting programmes and activities that have been developed to support students making AI programs, including AI websites (e.g. Google AIY and Machine Learning for Kids websites<sup>6</sup>), the Wolfram Language (Wolfram, 2017) and the eCraft2Learn project. In our recent work on eCraft2Learn (Kahn & Winters, 2018), one of the new block programming languages *Snap!* (Harvey and Möning, 2010) has been extended to provide an easy-to-use programming interface to powerful AI cloud services (as well as other AI functions within the browser). Furthermore, pre-trained machine learning models can be loaded, extended, and used in student projects. Young people can design and build impressive intelligent artefacts by composing and customizing components provided by world-class AI teams. Among the reasons for doing this are that students may become better motivated and empowered to produce very capable artefacts, while learning about cloud services, artificial intelligence, and other advanced technologies. These future developments can proceed iteratively with user testing to provide insights into the young people's perception of AI. These kinds of initiatives, together with a wider focus on media and digital literacy (e.g. Krutka et al., 2019; Pangrazio, 2016; *reference removed for blind review*) is an important component of change. With much discussion today focused on the risks of employment due to the automation of many jobs, ways in which young people, *from all backgrounds*, can be supported to program AI services promotes learning and understanding of AI's potential and the skills by which to use it in new, creative and effective ways.

Young people are, of course, not the only stakeholders who need to develop understandings about AI algorithms, teachers and other experts in education also need to learn about what data and AI is good for and what it is not so good at. Tackling the rhetoric around AI as the ultimate solver of societal problems to demonstrate the partial, often narrow, and general limitations of AI systems in education and society are important, to ensure educators expertise is still valued and made use of (Ozga, 2016; Ratner et al., 2018).

### **3.3 Working with the technical community to make AI systems accountable**

Digital structural violence has two core aspects which we now discuss in more detail: (1) its underlying technical process and (2) the associated strategies for their implementation. An example of a technical process is the composition of training data and the code developed to analyse it. An example of an implementation strategy is a procedure to ensure that the needs of marginalised learners are addressed.

As technical processes are often hidden and both AI and implementation are not well understood by the general public, transparency and accountability have become key concerns (Pasquale, 2015). From a sociological perspective, this is urgent work as understanding these processes is increasingly becoming a key form of human (and even social) capital. As Porter notes “the quantifiers have constructed the world in which they excel.” (Porter, 1994: 403). Thus, a key role the technology community can play is to make their data easier to engage with and their algorithms and decision processes easier to understand. This would enable educational experts of all kinds to utilise data in ways that were appropriate – including challenging and in some cases disregarding it. We already know that the opaque nature of the data, and the illusion of complete accuracy make it difficult for stakeholders to navigate the

---

<sup>6</sup> <https://aiyprojects.withgoogle.com> and <https://machinelearningforkids.co.uk>

differences between the data and their own expertise (Ozga, 2016) so how are they supposed to engage with more complex forms of data and machine learning?

Research over the last few years has focused on the need to make the decision-making processes assigned to AI algorithms more accountable, particularly within the Communication Studies community. Two positions can be delineated. In the first, Sandvig et al. (2016) argue that there is “new knowledge to be gained by considering the operation of algorithms writ small—the technical details of the innards of particular computer code” (p. 4974). They advocate for establishing multidisciplinary teams to develop our understanding of how AI algorithms make decisions, subjecting the underlying processes to ethical analysis. This is important, they say, because “we increasingly delegate authority to algorithms, and they are fast becoming obscure but important elements of social structure” (p. 4972).

### **3.3.1 Algorithmic transparency is not enough**

An alternative position is put forward by Ananny & Crawford (2016), who discuss the inadequacies of algorithmic transparency. Instead, drawing on research in experiential and constructionist learning, specifically the work of Resnick et al. (2000), they argue that learning about any complex system requires more than “being able to look inside systems or take them apart” (p.9). Instead, we need to be able to interact with systems in order to “understand how they behave in relation to their environments” (p.9). In line with Science and Technology Studies (STS) researchers, and what is already known about the privilege of experts and how social power is reconstituted through technical processes (outlined above), they advocate moving away from framing instances of AI in purely technological terms. This is not a new call. Our previous research has addressed it in the context of health care training [*removed for blind review*]. In the context of Education, Ball has long discussed the problems of prescriptive systems of accountability: “The question that is avoided in the heat and noise of performance management is whether these indicators actually ‘stand for’ and thus ‘represent’ valid, worthwhile or meaningful outputs” (Ball, 2017: 163).

Returning to the focus on the specifics of AI, Ananny and Crawford seek to overcome inadequacies of only focusing on algorithmic transparency by discussing the need to consider the wider complex system involving multiple actors, people environments, technologies etc., or what is termed an assemblage in Actor-Network Theory (for details on ANT, see: Callon, Law & Rip, 1986; Latour, 1987). They emphasise that “making one part of an algorithmic system visible—such as the algorithm, or even the underlying data—is not the same as holding the assemblage accountable” (Ananny & Crawford, 2016 p. 12).

The technical community have also been making strides to understand social issues. For example, there is an ACM Conference on Fairness, Accountability, and Transparency (ACM FAT\*) and an open access book in development on ‘Fairness and Machine Learning: Limitations and Opportunities’ was published at the time of writing (Barocas, Hardt & Narayanan, 2019). To illustrate the nature of this work, one tutorial given by Google AI at FAT\* 2019 focused on: “the quantitative definitions of fairness that arose in education and hiring in the wake of the U.S. Civil Rights Act of 1964, including considerations of the social, political and legal motivations behind them. We translate this quantitative research into modern day ML notation” (Hutchinson and Mitchell, 2019:np). While from a social science perspective there remains much work to do, it is nevertheless a positive development that computer scientists are engaging with these issues and, just as social scientists need support

from the AI community to learn about computer science, computer scientists should be supported in developing their understandings of STS. In support of this goal, STS researchers have been inputting to the FAT\* research agenda by pushing strongly for the social to be considered alongside the technical (e.g. Selbst et al., 2019) within this emerging field.

In summary, making the technical details of the computer code underlying AI algorithms available is laudable. As Burrell (2016) notes, algorithms need to be less obscure. Yet this is not sufficient. In many current cases, algorithms do not have a fixed observable state (although this may change with new research findings), and so information visualisations or other meaningful instantiations of the process may be difficult to produce. Moreover, making resources open source can (paradoxically) create greater inequality: as more and more resources are made freely available, those with existing access tend to benefit most. Those who did not have any access to begin with remain excluded unless additional steps are put in place to help them.

Reflecting on the suggestion of “holding the assemblage accountable” (Ananny and Crawford, 2016:12), may well be a preferable approach, but achieving it in practice is complicated. Seeking to make systems accountable through understanding the roles of multiple actors in an actor-network is beset with theoretical and methodological challenges. Conceptually, the tradition in ANT is that actants (i.e. things, people, objects, etc.) are not usually examined when they work and are understood as ‘black boxes’. They are opened up when they fail “to see how they should have worked and what went wrong” (Oliver, 2011 p. 380).

Working with the technical community to make AI systems accountable in Education will be a challenge. However, putting forward ways in which contextual and technical factors can be understood and accounted for is worth pursuing if we are to determine how specific forms of digital structural violence originate and are maintained.

#### **4. Conclusion**

As noted by Farmer (2004, p. 315), “Structural violence takes on new forms in every era”. What we do not want to see happen is AI embedded in the unseen structures of learning technologies to become a key component of a wider system of exclusion. It cannot be a mechanism of structural violence but as we have discussed, based on current projections, digital structural violence may well be an outcome of future learning systems. In highlighting this as an issue, what we have discussed is not really a grand conceptualisation, rather we have focused on introducing the concept and discussing how it may be addressed through future work in education and technology.

At present, much of the debates around AI and inequality are taking place outside the field of Education, and this is highly problematic given the emancipatory role that Education can play, and the multiple ways that AI is, and will be a feature of both informal and formal learning, across the life course. Where many commentators only focus on one of three proposed areas of focus: including marginalised learners in design, developing critical understanding through production, and working with the computer science and wider community to make these systems more accountable, we argue that all three of these strands of work need to be meaningfully connected together as part of an interdisciplinary research agenda.

In making this case, we are well aware of the challenges. Much has been written about the challenges of participatory design (Heeks, 1999; Bossen, 2006; Hirom et al., 2017) and there are questions about the extent to which educational stakeholders would realistically engage. For hard-pressed teachers, it is unrealistic to require them to develop in-depth algorithmic understanding; given the complex array of pressures they experience (Larke, 2019), yet finding ways to ‘open up’ algorithms to at least some level of inspection could be an important part of developing understandings in the ways discussed in section 3.2. Similarly, for young people, providing information on the data collected about their learning practices over time and detailing how this is used could be a useful starting point. This could go some way towards addressing the on-going concerns of researchers regarding the negative impact of data surveillance on learner agency (Selwyn, 2015). Perhaps of the three, the third strand (i.e. making these systems more accountable) is the most challenging, but would be made more possible if the first and second strand were also ongoing. Yet despite these challenges, we believe an interdisciplinary endeavour of the kind outline here is essential.

Our motivation with this paper has been to provide a concrete starting point for a material form of resistance to the reproduction of power structures. Through bringing together a conceptual framing, methodological approach, learning tools and supports for interdisciplinarity we can respond to the risk of a potential explosion of digital structural violence in future learning systems against the marginalised in the coming years.

## References

- Ananny, M. and Crawford, K. (2016) Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability, *new media & society*, 1-17.
- Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2016) Machine Bias: There's software used across the country to predict future criminals and it's biased against blacks. *ProPublica*, May 2016.
- Ball, S.J. (2017) *The education debate*. Policy Press.
- Barocas, S., Hardt, M. & Narayanan, A. (2019) Fairness and machine learning: Limitations and Opportunities. Available at: <https://fairmlbook.org> (open access)
- Bergviken Rensfeldt, A., Hillman, T., & Selwyn, N. (2018). Teachers ‘liking’ their work? Exploring the realities of teacher Facebook groups. *British Educational Research Journal*, 44(2), 230–250. <https://doi.org/10.1002/berj.3325>
- Bourdieu, P. (1977) *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.
- Bourgois, P. and Scheper-Hughes, N. (2004) Response to Farmer, P. (2004) An Anthropology of Structural Violence, *Current Anthropology*, 45(3), pp. 305-325
- Bossen, C. (2006) Participation, power, critique: constructing a standard for electronic patient records. In *Proceedings of the ninth conference on Participatory design: Expanding boundaries in design - Volume 1 (PDC '06)*, Vol. 1. ACM, New York, NY, USA, 95-104. DOI: <http://dx.doi.org/10.1145/1147261.1147276>

boyd, d. and Crawford, K. (2012) Critical Questions For Big Data, Information, *Communication & Society*, 15:5, 662-679

Buolamwini, J and Gebru, T. (2018) Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research Conference on Fairness, Accountability, and Transparency, 81: 1-15.

Burrell, J. (2016) How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*. <https://doi.org/10.1177/2053951715622512>

Caliskan, A., Bryson, J. and Narayanan, A. (2017) Semantics derived automatically from language corpora contain human-like biases, *SCIENCE*, 14 APR 2017 : 183-186

Callon, M., Law, J. and Rip, A. (1986). Mapping the dynamics of science and technology: sociology of science in the real world. Basingstoke: Macmillan.

Crawford, C., Dearden, L., Micklewright, J., and Vignoles, A. (2016). Family Background and University Success: Differences in Higher Education Access and Outcomes in England. Oxford: Oxford University Press

Daly, A., Devitt, S.K. and Mann, M., (Eds). 2019. Good data. Theory on Demand. Institute of Network Cultures, Amsterdam. Available from  
[https://eprints.qut.edu.au/125605/1/Good\\_Data\\_book.pdf](https://eprints.qut.edu.au/125605/1/Good_Data_book.pdf)

Daniels, H, Cole, T (2010) Exclusion from school: short-term setback or a long term of difficulties?, *European Journal of Special Needs Education*, 25(2), 115-130. DOI: <http://doi.org/10.1080/08856251003658652>

Davies, C., & Eynon, R. (2013) Teenagers and technology (pp. 1–163). London: Taylor and Francis.

D'Ignazio, C and Klein, L. (2019) Data Feminism. Cambridge MIT Press.

Drentea P and Moren-Cross J (2005) ‘Social capital and social support on the web: The case of an Internet mother site’. *Sociology of Health and Illness*, 27(7), pp. 920–943.

DuNann Winter, D. and Leighton, D. (2001). “Structural Violence: Introduction” in: D. Christie, R. Wagner and D. DuNann Winter (eds.), *Peace, Conflict, and Violence: Peace Psychology for the 21<sup>st</sup> Century*, Upper Saddle River: Prentice Hall, p. 99

Eubanks, V. (2012). *Digital Dead End: Fighting for social justice in the information age*. Cambridge: MIT Press.

Eubanks, V. (2018). *Automated Inequality*. St Martin’s Press.

Eynon, R. (2013) The rise of Big Data: what does it mean for education, technology, and media research?, *Learning, Media and Technology*, 38:3, 237-240, DOI: 10.1080/17439884.2013.771783

Eynon, R. (2015) The quantified self for learning: critical questions for education, *Learning, Media and Technology*, 40:4, 407-411, DOI: [10.1080/17439884.2015.1100797](https://doi.org/10.1080/17439884.2015.1100797)

Eynon, R. (2018) Feminist perspectives on learning, media and technology: recognition and future contributions. *Learning, Media and Technology*, 43(1).

Eynon, R. and Geniets, A. (2016) The digital skills paradox: how do digitally excluded youth develop skills to use the internet?. *Learning, Media and Technology*, 41(3), pp.463-479.

Farmer, P. (2004) An Anthropology of Structural Violence, *Current Anthropology*, 45(3), pp. 305-325

Funes, M & Mackness, J (2018) When inclusion excludes: a counter narrative of open online education, *Learning, Media and Technology*, 43:2, 119-138, DOI: [10.1080/17439884.2018.1444638](https://doi.org/10.1080/17439884.2018.1444638)

Galtung, J. (1969) Violence, Peace, and Peace Research, *Journal of Peace Research*, 6:3,167–191.

Green, B. (2018). Data Science as Political Action: Grounding Data Science in a Politics of Justice. *arXiv preprint arXiv:1811.03435*.

Hansen, J. and Reich, J. (2015) Socioeconomic status and MOOC enrollment: enriching demographic information with external datasets. In Proceedings of the Fifth International Conference on Learning Analytics And Knowledge (LAK '15). ACM, New York, NY, USA, 59-63. <http://dx.doi.org/10.1145/2723576.2723615>

Harding, S. and Hintikka M. Eds. (1983). *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*. Dordrecht: Reidel

Hargittai E., ‘Digital Na(t)ives? Variation in Internet Skills and Uses Among Members of the ‘Net Generation.’ *Sociological Inquiry*, 80, 2010: p. 92–113/

Hartsock, N. (1983). “The Feminist Standpoint: Developing the Ground for a Specifically Feminist Historical Materialism” in: Harding & Hintikka, (Eds), *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*, p. 284cc.

Harvey, B., Mönig, J., (2010) Bringing “No Ceiling” to Scratch: Can One Language Serve Kids and Computer Scientists? In Proceedings: Constructionism, Paris, France.

Heeks, R. (1999) The tyranny of participation in Information Systems: Learning from Development Projects. In Development Informatics Working Paper 4. Manchester, UK: Development Informatics Group, University of Manchester.

Hirom U., V.S. S., Doke P., Lobo S., Devkar S., Pandey N. (2017) A Critique on Participatory Design in Developmental Context: A Case Study. In: Rau PL. (eds) Cross-Cultural Design. CCD 2017. Lecture Notes in Computer Science, vol 10281. Springer, Cham

Hutchinson, B. and Mitchell, M. (2019) Translation Tutorial: A History of Quantitative Fairness in Testing, Tutorial given at FAT\*2019, January 2019, Atlanta, GA USA

Kahn, Ken (1977) Three Interactions between AI and Education. Machine Intelligence 8.

[Reference removed for review]

Kay, M., Matuszek, C. and Munson, S. (2015) Unequal representation and gender stereotypes in image search results for occupations. In, Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 3819-3828.

Kenway, J., Bigum, C., Fitzclarence, L., Collier, J. & Tregenza, K. (1994): New education in new times, *Journal of Education Policy*, 9:4, 317-333.  
<http://dx.doi.org/10.1080/0268093940090403>

Kidd, C.D., Orr, R., Abowd, G.D., Atkeson, C.G., Essa, I.A., MacIntyre, B., Mynatt, E., Starner, T.E. and Newstetter, W. (1999) The Aware Home: A Living Laboratory for Ubiquitous Computing Research. In *Proceedings of the Second International Workshop on Cooperative Buildings — CoBuild'99*.

Krutka, D. Manca, S., Galvin, S., Greenhow, C., Koehler, M., and Askari, E. (2019). Teaching "Against" Social Media: Confronting Problems of Profit in the Curriculum. Teachers College Record. 121: 1-19

Larke, L. (2019) Agentic neglect: teachers as gatekeepers of England's national computing curriculum, *British Journal of Educational Technology*, (Early view).  
<https://doi.org/10.1111/bjet.12744>

Latour, B. (1987) Science in Action: How to Follow Scientists and Engineers Through Society. Milton Keynes: Open University Press.

Loveday, V. (2018). The neurotic academic: anxiety, casualisation, and governance in the neoliberalising university. *Journal of Cultural Economy*, 11(2), 154–166.  
<https://doi.org/10.1080/17530350.2018.1426032>

Luckin, R., Holmes, W., Griffiths, M. & Forcier, L. B. (2016). Intelligence Unleashed. An argument for AI in Education. London: Pearson.

Macgilchrist, F. (2019) Cruel optimism in edtech: when the digital data practices of educational technology providers inadvertently hinder educational equity, *Learning, Media and Technology*, 44:1, 77-86, DOI: [10.1080/17439884.2018.1556217](https://doi.org/10.1080/17439884.2018.1556217)

Manovich, L. (2011) Trending: The Promises and the Challenges of Big Social Data. In *Debates in the Digital Humanities*, edited by M. Gold, 460–475. Minneapolis, MN: University of Minnesota Press.

Miller, E. (2017). ‘Teachers on Twitter: Why you should join and how to get started.’ *The Guardian*. <https://www.theguardian.com/teacher-network/2017/apr/20/teachers-on-twitter-why-join-get-started-social-media>

Narayan, U. (1988). Working together across differences: some considerations on emotions and political practice. *Hypatia* 3(2): 31-47.

Noble, S.U., 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.

Oliver, M. (2011) Technological determinism in educational technology research: some alternative ways of thinking about the relationship between learning and technology, *Journal of Computer Assisted Learning*, 27, p.373–384

O’Neil, C. (2016). *Weapons of Math Destruction*. New York: Penguin Random House

Otto HU, Kutscher N, Klein A, and Iske S (2005) Social Inequality in the Virtual Space: How Do Young People Use the Internet? Results from Empirical Research about Online Use Differences and Acquiring Patterns of Young People. Available at:  
<http://www.kib-bielefeld.de/veroeffentlichungen.htm>

O’Sullivan, K., Robson, J., Winters, N. (2018) “I feel like I have a disadvantage”: How Socio-Economically Disadvantaged Students Make the Decision to Study at a Prestigious University. *Studies in Higher Education* <https://doi.org/10.1080/03075079.2018.1460591>

Ozga, J. (2016) Trust in numbers? Digital education governance and the inspection process. *European Educational Research Journal*, 15(1), pp.69-81.

Pane, Debra Mayes & Rocco, Tonette S. (2009). Critical Microethnography: The Search for Emancipatory Methods [101 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 10(2), Art. 12, <http://nbn-resolving.de/urn:nbn:de:0114-fqs0902129>.

Pangrazio, L. (2016). Reconceptualising critical digital literacy, *Discourse: Studies in the Cultural Politics of Education*, 37:2, 163-174, DOI: 10.1080/01596306.2014.942836

Pasquale F (2015) *The Black Box Society*. Harvard: Harvard University Press.

Pew Research Center (2018) “Many Turn to YouTube for Children’s Content, News, How-To Lessons”. Last accessed 10 Feb 2019. <http://www.pewinternet.org/2018/11/07/many-turn-to-youtube-for-childrens-content-news-how-to-lessons/>

Porter, T. (1994) Making Things Quantitative, *Science in Context*, 7 (3): 389–407.

Ratner, H., Andersen, B.L. and Madsen, S.R., (2018). Configuring the teacher as data user: public-private sector mediations of national test data. *Learning, Media and Technology*, pp.1-14.

- Reich, J. and Ruipérez-Valiente, J. (2019) The MOOC Pivot, *Science*, 11 Jan 2019: 130-131
- Resnick, M., Berg, R. and Eisenberg, M. (2000) Beyond black boxes: bringing transparency and aesthetics back to scientific investigation, *Journal of the Learning Sciences*, 9(1): 7–30.
- Rice, R. and C. Haythornthwaite (2006) ‘Perspectives on Internet Use: Access, Involvement and Interaction’, in L. Lievrouw and S. Livingstone (eds) *Handbook of New Media: Social Shaping and Consequences of ICTs* (2nd edn), pp. 92–113. London: SAGE.
- Rieder, B., Matamoros-Fernández, A. and Coromina, O. (2018) From ranking algorithms to ‘ranking cultures’: Investigating the modulation of visibility in YouTube search results, *Convergence: The International Journal of Research into New Media Technologies*, Vol. 24(1) 50–68 DOI: 10.1177/1354856517736982
- Robinson, L. (2009) A Taste for the Necessary: A Bourdieuan Approach to Digital Inequality. *Information, Communication & Society* 12 (4): 488–507
- Robson, J., Randhawa, A., & Keep, E. (2018) Employability Skills Development in Studio Schools. Report for the Edge Foundation and Studio Schools Trust. London: The Edge Foundation <http://www.edge.co.uk/research/research-reports/report-on-the-create-curriculum-framework-in-studio-schools>
- Robson, J. (2018). Performance, structure and ideal identity: Reconceptualising teachers’ engagement in online social spaces. *British Journal of Educational Technology*, 49(3), 439-450. <https://doi.org/10.1111/bjet.12551>
- Robson, J. (2016). Engagement in structured social space: an investigation of teachers’ online peer-to-peer interaction. *Learning, Media and Technology*, 41(1), 119-139. <https://doi.org/10.1080/17439884.2015.1102743>
- Sandvig, C., Hamilton, K., Karahalios, K. and Langbort, C. (2016) When the Algorithm Itself Is a Racist: Diagnosing Ethical Harm in the Basic Components of Software, *International Journal of Communication*, 10: 4972–4990
- Selbst, A., boyd, d., Friedler, S., Venkatasubramanian, S. and Vertesi, J. (2019) Fairness and Abstraction in Sociotechnical Systems, In *Proceedings of FAT\*2019*, January 2019, Atlanta, GA USA
- Selwyn, N. (2015) Data entry: towards the critical study of digital data and education, *Learning, Media and Technology*, 40:1, 64-82
- Tikly, L. and Barrett, A. (2011). Social justice, capabilities and the quality of education in low income countries. *International Journal of Educational Development*, 31, pp. 3–14

Vandewaetere, M., Desmet, P. and Clarebout, G. (2011) The contribution of learner characteristics in the development of computer-based adaptive learning environments, *Computers in Human Behavior*, 27(1), 118-130, <https://doi.org/10.1016/j.chb.2010.07.038>.

Weizenbaum, J. (1972) On the Impact of the Computer on Society. *Science* 176 (4035), p. 609-614.

Wessels, B., 2015. Digital Divide| Authentication, Status, and Power in a Digitally Organized Society. *International Journal of Communication*, 9, p.18.

White GW, Suchowierska M, Campbell M. Developing and systematically implementing Participatory Action Research. *Arch Phys Med Rehabil* 2004;85(4 Suppl 2): S3-12.

White, P. and Selwyn, N. (2012) Learning online? Educational internet use and participation in adult learning, 2002 to 2010. *Educational Review* 64(4):451-469.

Williamson, B. (2018). The hidden architecture of higher education : building a big data infrastructure for the ‘ smarter university ,’ 1–26. <https://doi.org/10.1186/s41239-018-0094-1>

[removed for blind review]

Wolfram, S. (2017) An Elementary Introduction to the Wolfram Language, Second Edition. Wolfram Media

Zuboff, S (2019) The Age of Surveillance Capitalism. London: Profile Books