Datafication and the role of schooling: challenging the status quo

[This is a draft version of a chapter. The reference for the final version is as follows:

Eynon, Rebecca. 2022. "Datafication and the Role of Schooling: Challenging the Status Quo." In *Learning to Live with Datafication: Educational Case Studies and Initiatives from Across the World*, edited by Luci Pangrazio, and Julian Sefton-Green, 17–34. London: Routledge.]

The increasing use of digital trace data has significant implications for understanding the purposes of education. This chapter begins by provides an overview of some of the key features of the datafication of education. It then examines how the use of digital trace data is influencing the central roles of schools, drawing on the work of Gert Biesta; and highlights how the current use of such data is likely to compound existing inequalities. The chapter then provides a brief account of current responses to digital trace data both within and outside education and argues that these are inadequate. The final part of the chapter argues for the need to reconfigure data use in schools via digital literacy and participatory design to help to challenge the status quo.

The datafication of education

There is a long history of data collection in education, and investment, interest and excitement about the potential of data grows year on year (Watters, 2014). At present, it is helpful to consider two broad kinds of digital data collected about young people and their education. The first consists of information about student demographics, course enrolment, and an array of summative and formative assessment data that schools are increasingly required to collect to meet the accountability demands placed upon them by governments. The vast majority of this data is collected via examinations, audit and accounting procedures, responses to surveys and audits, and large-scale government data gathering activities. Examples of such data include the OECD's Programme for International Student Assessment, and well-established national programmes such as the National Pupil Database (NPD) in the UK.

Over the past decade or so, there has been increasing attention on a second, 'new' kind of data, often described as Big Data or, more accurately, digital trace data. In essence, this is the petabytes of very fine-grained, real time data generated when an individual uses digital technology for learning. Systems that produce such data include learning management systems, digital games, augmented and virtual reality systems, intelligent tutors, adaptive assessments, social networking platforms, wikis, search engines, and blogs; and are encountered via a growing range of networked wearable, personal and mobile devices (Romero & Ventura, 2020; Luckin et al, 2016; Fischer et al., 2020). These interactions can take place in what are sometimes described as blended learning contexts, i.e. as a part of teaching that takes place face-to-face, or can be a primarily digital experience in or out of school.

Characteristics of digital trace data

Both kinds of data are in some sense big. Datasets of each type often contain millions of data points and each have implications for educational policy and practice. However, some characteristics of digital trace data are distinct from more traditional forms of educational data. The most well cited of these are the 3Vs – "Huge in *volume*, consisting of terabytes or petabytes of data; High in *velocity*,

being created in or near real time; Diverse in *variety* in type, being structured and unstructured in nature, and often temporally and spatially referenced" (Kitchin, 2014: 67, drawing on Laney, 2001).

Due to these characteristics, digital trace data offers a significant array of real-time, fine-grained data in ways that are not typically available with other methods. In relation to learning, for example, interactions between teachers, students and their peers, patterns of engagement with course material, learners' emotional and physical states can all be captured at a level of detail that may be difficult to observe in the classroom (Luckin et al., 2016; Fischer et al., 2020; Romero & Ventura, 2020). These insights are then used in schools to detect, understand and intervene in learning in some way. In some cases, such data might be used as part of an automated teaching and learning system (e.g. intelligent tutoring systems or adaptive assessments) or the basis of systems to inform students or teachers about progress (e.g. via LMS dashboards) that can aid education decision making. These are sometimes retrospective (e.g. mapping activities over the term) but are increasingly part of predictive systems (e.g. likely exam performance, and drop-out risk) with some systems also offering specific interventions to address these issues (Fischer et al., 2020; Hakimi et al., 2021).

Though learning is a core focus, there are a growing number of broader uses of digital trace data relevant to research and policy such as informing resource allocation, evaluating interventions, timetabling and accountability procedures. Indeed, such data can be used by individuals, schools, educational policy makers and the commercial sector in myriad ways, and can sit alongside and sometimes be combined with existing data driven processes such as school inspections; school league tables and policy comparisons (Ozga, 2016; Williamson, 2021).

Precisely what this kind of data offers and does not offer is the subject of much academic debate. Yet in policy and commercial discourse, digital trace data is typically framed as having almost magical properties, despite the realities being far more complex. Debates around Big Data are having a significant impact on understandings of what matters in social life and the knowledge claims society values. As Markham notes, Big Data "functions as a powerful frame for discourse about knowledge — both where it comes from and how it is derived; privileges certain ways of knowing over others; and through its ambiguity, can foster a self–perpetuating sensibility that it is incontrovertible, something to question the meaning of, or the veracity of, but not the existence of" (Markham, 2013 n.p).

This is clearly apparent in education systems where digital trace data is positioned as the solution for all current challenges in education. It is, in this sense, the latest in a long line of 'technical fixes' for education where fixing education is primarily about making the delivery of education more efficient and effective (Robins and Webster, 1989). As Ozga notes, for decades "there is a normalisation of digital data work within education / schooling, where it is conventionally understood and defended as the basis of improvement" (Ozga, 2016: 70). Digital trace data has added to this trajectory. Within this framing Big Data is positioned as having some kind of transformative power despite there being very little clarity about what precisely in education needs to be transformed or why such data is the answer to such problems.

This way of framing digital trace data is largely due to the efforts of the commercial sector. The role of the commercial sector in education, and the promotion of different kinds of data to support effectiveness and efficiency measures is not new. However, it is perhaps intensifying with the move to focus on digital trace data and related debates around Artificial Intelligence (AI) (Davies et al., 2020) where digital trace data is being used to make education function more like markets, i.e., in making education more accountable, measurable and comparable. Importantly, the use of digital

trace data was first used and promoted by the commercial sector to enhance profit and make efficiency gains; and these neo-liberal logics can be seen in many spheres of public life including education (Crawford et al., 2014).

The role of the commercial sector

As with any apparently new intervention, the ways that this new kind of data is being taken up, thought about and used in education has been significantly shaped by a complex network of actors that includes multinational businesses, the world bank, OECD; regional states and organizations like the EU and global philanthropies alongside national and local actors (Ball, 2017). A significant part of these networks are made up of commercial organisations who are influencing policy-making in ways that are likely to be beneficial to them.

These actors, then, offer both discursive and practical ways to influence policies in nation states, while also encouraging a more global agreement on the need for reform in education institutions (Ball, 2017). For example, Pearson who openly seeks to be "an active participant in national educational policy conversations" (Ball, 2017: 44), produces a series of software and applications available to schools; and also provide analytics insights to governments (Williamson, 2021). Companies such as Pearson not only offer the specific systems for schools, but also "own the algorithms and analytics required to make sense of those data" (Williamson, 2017: 105).

The growing prominence and power of the commercial sector has significant implications for education, as certain kinds of values are promoted, typically that support commercial agendas but also that assume a strong economic purpose for Education (Ball, 2017). Furthermore, through the increasing power of these actors, questions of education are no longer focused on the purposes or role of schools. Instead they have been replaced with questions simply about the best process of delivering education through the use of advanced data analytics (Biesta, 2015). However, such data is of course, not neutral and its use has implications for what schools 'do' and what their relationship is to society.

The purposes of schooling in an era of datafication

A strong and growing body of work has highlighted a number of problematic implications of the datafication of education. These include issues of privacy, surveillance, performativity, instrumentalism and governance (e.g. Porter, 1994; Ball, 2017; Williamson, 2021; Jarke and Breiter, 2019). Digital trace data intensifies these and other concerns in part due to its size, velocity, temporality and relational nature (Kitchin, 2014). The implications of digital trace data are not just a quantitative change (in terms of amount and nature of data collected about education) it also has potentially different qualitative implications (Webster, 2014). As Porter notes, new forms of data "must ask not only about their validity but also about how the world might be changed by adopting new forms of quantification" (Porter, 1994: 404).

Empirical studies of how these data-intense systems are used and taken up by educational institutions, are, as illustrated in other chapters in this book, an essential component of better understanding the social and educational implications of digital trace data. This chapter aims to complement these empirical studies within the context of exploring questions about the purposes of education (see also Eynon, forthcoming).

There are, of course, multiple ways of conceptualising the purpose of education. Here the discussion uses Biesta's very helpful framework (Biesta, 2015). He suggests a focus on three interrelated purposes or roles of education: qualification, socialization and subjectification. Qualification functions provide young people, "with the knowledge, skills and understandings and often with the

dispositions and forms of judgement that allow them to 'do something'. (…) [This] can range from the very specific (such as in the case of training for a particular job or profession, or the training of particular skill or technique) to the much more general (such as an introduction to modern culture, or the teaching of life skills etc)" (Biesta, 2015: 19-20). Socialisation, in other words, the ways that "we become part of particular social, cultural, and political 'orders'" (ibid, p. 20) is another important function of schooling. This can be both intentional and unintentional and can be both positive and negative. In this way, schools are important in the "continuation of culture and tradition" (Biesta, 2015: 20). Subjectification, is "the process of becoming a subject" (ibid, p. 21) where the focus is on "the kinds of subjectivity that are made possible as a result of particular educational arrangements and configurations (….) that allow those educated to become more autonomous and independent in their thinking and acting" (Biesta, 2015: 21).

Biesta argues that reflecting on the optimum combination of the three roles of education provides a useful way to promote conversations about the kinds of education a society wants at different times and contexts and moves society away from a focus on data and measurement (Biesta, 2015). How then could the use of digital trace data influence these roles of a school?

Data and qualification

All three purposes of schooling (qualification, socialization and subjectification) are clearly important roles of education, yet in many countries there is a tendency to focus increasingly on the economic aspects of qualification (Biesta, 2015). Positioning schools in a way that prioritises societal needs, particularly those of the economy, has narrowed perspectives about what schools are for, from providing a democratic education towards an economic agenda of developing skills for economic gain in the job market (Powers et al., 2016). Educational outcomes (and indeed the curriculum as a whole) have become relatively narrow as the focus is on the economic needs of society (Biesta, 2015).

This trend of narrowing outcomes is exacerbated by the increasing use of digital trace data due to: 1) the ways it can narrow the curriculum through measurement; 2) change the content of the curriculum; and 3) aid in standardizing the curriculum.

Measurement is a process, where we "cut up the world" into discrete entities. For any computational analysis to take place, "the continuous flow of our everyday reality [is transformed] into a grid of numbers that can be stored as a representation of reality" (Berry, 2011: 1-2). Data does not straightforwardly represent reality (Borgman, 2015; Espeland and Stevens, 2008). Data are socially constructed representations that are shaped by the expertise and background of the person or people engaged in data collection and analysis, their current motivations for using the data, shaped by temporal and geographical circumstances (Borgman, 2015; Boellstorff, 2013).

What is assessed in education has long been a contested topic, both in terms of the validity of what is measured, and what counts as educationally meaningful (Ball, 2017). With the increasing use of novel forms of trace data in education, data scientists are making these choices to determine which measures to include to best fit to a model and what the most appropriate outcome measures should be. Indeed, sometimes such choices are made by the algorithm as opposed to a human. In both cases, it is not always clear what measures in data models really represent, even if they improve model fit. Similarly, the outcomes of such models are typically very narrow to capture a complex aspect of schooling but such assessments were only ever intended to be a proxy for wider achievements (e.g. Luke 2009). Nevertheless, the measures then become encoded as the central goals of education and prioritize aspects of education that can be measured easily rather than aspects of education that society values but are far more difficult to measure (Biesta, 2015).

The designers of data intense systems for education have an increasing opportunity to shape the curriculum in ways they see as appropriate because they are building the system. As the majority of such systems are built by the commercial sector their content is based on topics that the data and EdTech companies think matter. As we have seen throughout educational history, it is not just *what* knowledge is most valued, but *whose* knowledge is most valued when thinking through questions of power in education (Apple, 2012). Data scientists and the commercial sector have more power than ever in determining what kinds of knowledge, whose histories, and whose content counts. This is not only about narrowing the curriculum, but also standardizing it. Monolithic approaches or scaled interventions are something education technology companies desire as this way they can maximise profit and reduce costs. Such activities, automate and standardize knowledge, curriculum and pedagogy in ways that are highly problematic (Saltman, 2016).

The increasing focus on a narrow set of scores has multiple impacts on education, often as they are not only used to test students, but as part of accountability mechanisms for teachers and schools. This leads to risks of teaching to the test (Selwyn, 2019), impacting pedagogical choices of teachers both in terms of how they teach and also what they teach (Roberts-Holmes and Bradbury, 2016) and encourages performativity and learning for the test in students (Roberts-Holmes and Bradbury, 2016). These challenges are not new, but data driven systems are reinforcing and intensifying the long-entrenched emphasis on education as an exam-oriented activity and closing down any possibility of moving away from that negative kind of approach to education. Importantly, these trends and concerns with particular parts of the curriculum or particular kinds of knowledge have wider implications than what students learn. As Biesta notes, "students do not just learn from the content we provide them with but also from the ways in which we provide them with this content" (Biesta, 2019: 4) and this has implications for socialisation and subjectivity.

Data and socialisation

Socialisation draws attention to the ways that young people are socialised into existing orders, the ways that schools "(re)present cultures, traditions, and practices, either explicitly but often also implicitly" (Biesta, 2015:20).

Data can shape how individuals think about these cultures, traditions, and practices because it encodes and shapes how actors think about themselves and their relationships to others. Well established and important examples include the case of gender or intelligence, where data is used to define neat categories and boundaries from messy and complex realities (Espeland and Stevens, 2008). In doing so it shapes people's expectations about the world and their identity within it, it assists in a process 'making up people' (Espeland and Stevens, 2008, drawing on Hacking, 1999). This process influences everyone, but is particularly pertinent for young people who are determining who they are in the world (Davies and Eynon, 2013).

When teaching and learning is to some degree automated, as is typically the case with the use of digital trace data, it is likely to have effects on how students think about themselves and the expectations society has for them (Apple, 2012; Saltman, 2016). Data intense systems in education can encode expectations of what a learner should be and how they should act (Decuypere, 2019) or indicate what the future holds for young people like them (Eynon, forthcoming). A concern raised by many critical scholars is the ways that digital trace data is being used in schools supports a strong measurement regime, where constant monitoring becomes the norm (Jarke and Berieter, 2019) which is likely to lead, among other things to an acceptance of monitoring in all contexts and life stages. An important question, is what kind of world, characterised by the use of digital trace data, does society want young people to be socialised into?

Data and subjectification

The role of data driven systems in schools are often seen as a way to free up the time of teachers. This argument is not new – Skinner, one of early pioneers of teaching machines also argued for the importance of using such machines to free up teacher time (Skinner, 1961).

Yet, automating parts of the teaching process changes how learning and teaching happens. Importantly, they "displace the dialogue between teachers and students with prepackaged curricula" (Saltman, 2016: 113); and encourage more didactic forms of learning new things as opposed to encouraging pedagogical exchange that is important for education (Saltman, 2016: 108). The relationships between teacher and learner are an important part of education, and changing it has consequences. As Audrey Watters asked in a keynote at Berkeley City College in 2014, "what does it mean to tell our students that we're actually not going to read their papers, but we're going to scan them and a computer will analyze them instead?" (Watters, 2014: 56).

Even if data intense systems in schools lead to students scoring more highly on certain tests, it is important also to think more about whether this is educationally desirable (Biesta, 2015). Although many academics and teachers are concerned about the extent to which such systems may have for depth of understanding, these systems may also undermine subjectification, that is, people's capacity for action (Biesta, 2015). This 'coming into presence' (Biesta, 2016) is only possible when others acknowledge and respond to our actions (Licenberg and Eynon, forthcoming). If students feel that their actions are not meaningfully engaged with by teachers or by other students, as could be the case in some systems, this has significant implications for education. The behavioural engineering practices such as nudging to produce learning outcomes that are characteristic of many of these systems (Knox et al., 2019) make it difficult for students to "respond in their own unique way to the learning opportunities provided by the curriculum" (Biesta, 2016: 138) in a way that facilitates a genuine interaction between teachers and students (Licenberg and Eynon, forthcoming).

Furthermore, these nudging characteristics of such systems might disempower and infantilize students (Selwyn, 2019). As Hartman-Caverly (2019) suggests, an individual's ability to determine what and how to deploy attention is central to the development of people as "self-sufficient learners and independent thinkers" (Hartman-Caverly, 2019: 24). The use of data methods in education such as nudging or trying to shape learner's behaviour in any way can interrupt this process, directing the attention of students (preventing intellectual freedom) and disrupting our attention and the possibilities of focusing on an issue in an in-depth way (Hartman-Caverly, 2019). This is not just about the risks of the negative impacts of data driven systems on the possibilities for sustained engagement for meaningful learning, but the impact this may be having on our abilities to become a person who can think and act independently and to learn in a self-determined way (Hartman-Caverly, 2019).

Compounding inequalities

The influence of digital trace data on the roles of schools will vary. Thinking through the implications of using digital trace data using the framing proposed by Biesta varies significantly according to context. This chapter is written primarily from the perspective of schooling in the global north, and even within this collection of countries, experiences, practise and policies will vary significantly.

Schools are of course highly varied places. While they may experience similar pressures and constraints from policy makers and other systems, schools vary in their practices and philosophies and the ways that they engage with and react to policy and other pressures (Lipman, 2002). Teachers and other stakeholders all have some agency in this process, they can all, "'rewrite' policies

through their own actions within the restrictions imposed on them" (Lipman, 2002: 383). Nevertheless, some will have more opportunities to exert their agency than others.

Research that has explored digital inequalities in schools have long documented the differences in what some of those responses may look like (Warshauer, 2004); and it is likely the same variation will be seen with digital trace data. There are concerns that schools with less resources tend to resort to more standardised / assessment driven tools and in general the data tends to be more likely to dictate practice rather than be used as one of many approaches. As Zeide notes, "students enrolled in schools which can afford to allow teachers to deviate from algorithmic recommendations or supplement automated assessment with personal evaluation will have the opportunity for accommodation based on individual circumstances. Less fortunate students may receive automatically differentiated instruction without the flexibility of contextualized assessment" (Zeide, 2017:169). Indeed, outside of schools, forms of educational redlining are already emerging (e.g., the differential use of AI bots in MOOCs) with those able to pay provided with a more sophisticated human-mediated experience (Winters, et al., 2020).

This argument fits with wider concerns that draw attention to the need to focus on who is subjected to the analysis of Big Data systems, as those who are most marginalised and least well off are more likely to be targeted (Crawford et al., 2014: 1666). As well as perhaps being over-reliant on data systems as discussed above, the use of education digital trace data makes even more visible certain groups of students, "The 'outliers', 'under-achievers', and 'under-performers' produced by performance measures become targets of manipulation, disapproval and anxious self-scrutiny" (Espeland and Stevens, 2008:416, drawing on Hacking, 1999).

Studies have shown how students who are classified negatively in some way are then subject to different interventions. For example, in studies of primary school children in the UK children who were on a passing borderline were excluded from arts-based lessons to receive intensive phonics booster classes (Roberts-Holmes & Bradbury, 2016). Higher achievers were left to succeed on their own, a low achieving group were considered hopeless cases, and those children in the middle ground were given enough support to get them to the right level (Roberts-Holmes and Bradbury, 2016). As Spielman notes, "it is a risk to social mobility if pupils miss out on opportunities to study subjects and gain knowledge that could be valuable in subsequent stages of education or in later life. Restricted subject choice for low-attaining pupils disproportionately affects pupils from low income backgrounds" (Spielman, 2017: n.p.) yet this approach, due to the demands of league tables is routinely happening in some schools.

These trends are perhaps not new, but intensify with data driven systems, and indeed are likely to be further compounded by the likely biases in these systems. Systems designed to: facilitate school choice and integration, evaluate essay writing, detect concentration and emotion in the classroom, evaluate the effectiveness of teachers, assess students, check for cheating, and map attendance are just some of a growing number of examples of systems that can favour certain groups of students over others (Crawford et al., 2019; O'Neil,2016; Watters, 2020; Whittaker et al., 2018).

Current responses

Given the arguments made above, it would seem appropriate for a robust response to ensure the ways that digital trace data is used in schools leads to the kinds of education systems that society wants for young people. Yet, in reality the current responses both in and outside school are inadequate. This is, for the most part, because the practical and policy orientated responses to digital trace data essentially treat such data as if it is simply a bigger or better set of data than has

traditionally been collected in education systems. It is not considered as something that can lead to qualitative changes in the role of schooling, just as something that improves existing processes.

Outside education

This has implications for the kinds of policy responses to digital trace data that have been seen across the globe. For example, by framing digital trace data as a straightforwardly good thing, collecting student data becomes acceptable as more data leads to better models, and this will lead to digital trace data doing the best for the most students. Within such a frame, discussions move quickly to the legal basis to protect individuals' information or other appropriate governance practices. Thus issues of the use of digital trace data becomes a debate about legal and practical requirements rather than wider questions of values within education (Bulger, 2016; Eynon, 2013; Hakimi et al., 2021). Issues of privacy and data protection, for example, are extremely important but not sufficient.

Legal frameworks are problematic as they tend not to keep up to date with technical trends and also do not account for the wider social and educational implications of such approaches. For example, a legal analysis of US policy highlighted that legal provisions in schools focused on information practice principles, but this neglected more value orientated concerns around the use of such data such as questions of student agency or the risks of a discrimination (Regan and Jesse, 2019). Similarly, while there has been significant progress in relation to privacy and data protection (e.g. the EU's General Data Protection Regulation (GDPR), introduced in 2018) that is important in education, it does not fully account for the ways that the move towards data driven systems in education has implications for human rights (Berendt et al., 2020).

Alongside legal responses, many companies are setting up governance structures. Yet, there are concerns that these governance structures operate as self-regulatory checklists that substitute for independent legal or regulatory oversight (Wagner, 2018). Many of these governance structures are increasingly contested, including concerns about practices of 'ethics washing' where companies promote an image of concern about ethics, while fundamental practices remain unperturbed by these public facing activities (Wagner, 2018). They also tend to focus on security, transparency, accountability and reliability of data (Hakimi et al., 2021) – which while important does not fully address the kinds of issues raised above.

Relatedly there are growing concerns that the commercial sector has too much power in determining ethical governance and regulation of the use of digital trace data and AI across all areas of social life (Benkler, 2019). As Macgilchrist (2019) notes, "no matter how good the motives, and how pedagogically well-founded the decisions, it is a post-democratic moment when the ability to make these decisions has shifted from publicly accountable government officials, policy-makers or educators, to developers, programmers, designers and other staff in private edtech organisations" (2019: 83).

Within education

Schools, for the most part, have echoed this primarily instrumental response. In other words, they tend to support (in practical terms) the use of data to enhance efficiency, increase transparency, support competitiveness, evaluate performance and improve the learning experience. This fits with a wider concern around school policies in the recent times, that school policy is largely reactive instead of proactive, in that it is based on a relatively passive response to the perceived requirements of society. It is based on the assumption that there is a need to act and benefit from the changes occurring in our society, and to do so fast, rather than thinking about schools in a way that help create the society we want (Biesta, 2013).

This is likely partly due to the ways data encourages the creation of education markets that encourage "a culture of self-interest" by promoting parental and student choice (Ball, 2017: 54), thus focusing attention on individual and school needs, as opposed to wider social and educational concerns. Performativity, both of educators and of students, is a significant concern (Ball, 2017). Data encourages a situation where teachers begin to "'want' what the system needs in order to perform well" (Lyotard, 1984: 62)" (Ball, 2017: 55). For example, a study of primary schools in the UK highlighted how even though teachers were aware of the problems of all the data they were asked to collect, they tended to comply with the requests, adding more self-governance to an already challenging workload (Robert-Holmes and Bradbury, 2016). Schools stop talking about what schools are for, and instead focus on indicators of "quality [that] are taken as a definition of quality" (Biesta, 2019: 3).

The acceptance of data as a good thing is also reflected in the coding agendas seen in many school systems. Coding is promoted as a way to guarantee success for the individual and society – but coding does not translate into the same opportunities for everyone. All young people have different levels of economic, social, cultural and political resources and this clearly influences the kinds of opportunities they are able to get from new technologies. Yet, the focus has become very narrow (on code literacy) and not on data literacy (which is what would enable people to think about data structures as well as algorithms) (Driscoll, 2012). As a society, it is important to make decisions about how society wishes young people to use and engage with data-intense systems, and how education systems are developed to support that vision.

Reconfiguring data use in schools

Digital trace data has all kinds of implications for the roles of schools that require a considered response. This requires more meaningful debate amongst all actors in the ecosystem: policy makers, the commercial sector, academics, students and the general public. As part of this strategy, it is imperative to move away from a system where commercial interests dominate the discussion.

It is perhaps a good moment for such a shift. The Covid-19 pandemic has thrown light on the importance of schooling, and the complex role schools play in society; it has also made visible the significant social, educational and technological inequalities in society. The examination controversies in the summer of 2020, such as those led by the UK government and Ofqual, made visible the significant problems of current assessment systems and raised questions about the extent to which data and algorithms should be privileged over professional judgement. At the same time, teachers and students were often given little choice but to learn and teach via digital technologies, and have found the many limitations as well as possibilities that these data intense systems offer.

At a time when many are looking for positives in the most difficult of times, with hopes for creating the 'new normal' and 'building back better' there may be opportunities to raise debate in this area. For example, the experiences during the pandemic may well enable broader conversations about the role of school in societies and how technology and related data driven systems should be used in Education. Teachers may have a new found respect in society and a return to a recognition of their professional status, with more power and a stronger sense of agency in discussions over the current and future use of data systems in education.

There is then, the potential of a heightened awareness of data-related issues amongst educational professionals and the public, which may enable sustained debate over what ends societies want data-driven education to achieve (Selwyn, 2019). Such dialogue is difficult to enable but is possible, particularly when facilitated at a local level with local concerns in mind. For example, drawing on the case of the Citizen Schools of Porto Alegre, Apple (2012) documents how "a wider group of

people (including local citizens, teachers, administrators, parents) [were] actively involved in decision making – about budgets, about supporting knowledge construction that [was] centred within the interests and concerns of the local community as well as connecting to the national curriculum" (Apple 2012:138). Despite this group of people not ordinarily "allowed" or expected to be part of the discussions to help to enable change, positive change occurred, with participatory mechanisms at the heart of the strategy (Apple, 2012:103).

Relatedly, D'Ignazio and Klein (2019) draw on feminist theory to argue for the importance of multiple actors needing to be involved in data work. Their approach 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. 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.). This could be highly promising when applied to educational settings.

One key approach of more participatory methods is participatory design. Participatory design has a long history of enabling both theoretical and practical change, and a number of authors have proposed its importance in enabling change in this area. Key to this is to actively design in the values and aims of Education that society wishes to promote, and in doing so may help to challenge and reinvent the current commercial eco-system (Macgilchrist et al., 2020). Relatedly, authors have called for an 'ethics by design' approach in educational institutions that use data analytics where ethical considerations and their renegotiation are at the heart of the design of data based education systems and involve key stakeholders at all stages (Gray and Boling, 2016). One option, for example, may be to use the framing proposed by Biesta (2015) to help negotiate and determine those values as part of a design strategy.

It is important to be mindful that participatory design, as with all participatory methods are not straightforward. Technologies can still be designed in ways that are problematic or reinforce inequalities and there is nothing to say that users are somehow expert in what is needed (Berg, 1998). Such techniques cannot be considered a straightforward solution, but they are likely to enable better theorisation, along with a richer understanding of key issues, and enable more conversations around data and education to occur. Indeed, there are some excellent examples in this book.

A second important area is ensuring students have control and understanding of their data. As Audrey Watters argued at a keynote in 2014, when talking about the metaphors and implications of data mining, "I want to encourage the building of technologies that see students' lives and learning not as a resource to be extracted but as something they themselves can control and cultivate" (Watters, 2014: 105). This is not just about individuals being more aware of and having more control over their data usage, but instead being part of something that enables them to create a new kind of public engagement, discourse and change around the ways that data in education are envisaged and used.

Similar to Couldry's (2003) discussion on the need to shift static discussions of a digital divide towards a discussion about how the Internet can be an active space that contributes to democratic life, it is not just whether people know how to access and use their data but if they can "effectively

represent their material interests" (Couldry, 2003: 94). The Internet, or indeed digital trace data, is not just something that exists and can be black-boxed, but "should remain, open to deliberative intervention and ethical inquiry" (Couldry, 2003: 90).

Education is of course a central part of this, and connects closely with calls for digital data literacy, where what is needed is a, "deliberately political model of digital literacy in which complex and detailed understandings of discourse, ideology and power in the digital context are scaffolded" (Pangrazio, 2016: 170). An important part of learning this is through the process of design, and enabling young people to produce their own artefacts (Pangrazio, 2016; Winters et al., 2020). Precisely how this can be achieved in multiple contexts is documented in other chapters throughout this book.

Young people are, of course, not the only stakeholders who need to develop understandings about digital trace data. Teachers and education stakeholders also need to learn about what data and AI is good for and what it is not so good at (Winters et al., 2020), but it is an important start and may help to reconfigure the existing direction of inquiry.

Despite the possibilities of using such approaches to reconfigure the status quo, it is important to note that schools do not operate in a vacuum. As Reay notes, "educational systems are only as good as the societies they emerge out of (…) capitalist, neoliberal societies beget capitalist neoliberal educational systems" (Reay, 2011:2). While data are at the heart of the business model of many other mainstream platforms that young people may use for learning and everyday life (Zuboff, 2019) there are significant implications for education (Hakimi et al., 2021). As Hartman-Caverly notes, "learning analytics is but one node in a broader network of surveillance capitalism, in which the power to accumulate information harvested from human behaviour creates control over others' lived realities" (Hartman-Caverly, 2019:40; drawing on West, 2019; Zuboff, 2015). Given the current direction of surveillance capitalism, there are important questions about how we educate for democracy in such a context (Biesta, 2019, Sefton-Green, 2020).

Conclusion: challenging the status quo

In summary, data is becoming a defining feature of our education system. Importantly, digital trace data or Big Data needs to be understood both technically but also socially and culturally. At present digital trace data is being framed in a particular way that promotes the marketization of schooling and as a way to solve all educational changes. Yet who benefits most from this framing are primarily those from the commercial sector, not from education.

Using Biesta's (2015) framing to critically examine the roles of schools highlighted how the use of digital trace data has multiple implications for what schools do. Such changes can be seen at all levels, from the experiences of individual learners to whole school systems – and is likely to further disadvantage those who are already disadvantaged.

If society wishes to continue using such data within schools it is important to reconfigure the ways it is being conceptualised, designed, and implemented, and this requires changes across all actors in the ecosystem. Schools should not just reflect society, they can also change it (Apple, 2012). Rather than merely reproducing and reinforce the dominant ideologies of those in power, there are possibilities for schools to instead question those patterns or indeed, facilitate a change in wider society (Apple, 2012). This can be achieved through various approaches including participatory methods and changes in curriculum from moving from a focus on coding to data literacy.

For centuries, there has been a debate about the most appropriate relationship between school and society. Are schools simply there to meet the needs and demands of society, or should they be

viewed as institutions slightly set apart from society? The first perspective emphasizes that schools provide whatever society needs at any given point in history, in the latter; schools are places of reflection and action that may challenge and change the status quo (Biesta, 2019). Biesta (2019) notes that both are likely necessary for any democratic society, yet that at present the balance between these two positions has been lost: the dominant position is that schools should service the needs of society.

The current faith in data and data solutionism, education policies driven by the data economy, and educational tools based on a philosophy of extracting economic value from the collection and processing of data resulting from digital technologies and the people who use them, where commercial interests are prioritised (Selwyn, 2019), does not provide many opportunities for change. Yet it is possible. The other chapters in this book will discuss in more depth how this could be achieved, and how schools can challenge the current status quo.

References

Apple, M.W. 2012. Can Education Change Society? London: Routledge.

Ball, S.J. 2017. The education debate. Policy Press.

Benkler, Y., 2019. Don't let industry write the rules for AI. Nature, 569(7754), pp.161-162.

Berendt, B., Littlejohn, A. and Blakemore, M., 2020. AI in education: learner choice and fundamental rights. Learning, Media and Technology, 45(3), pp.312-324.

Berg, M. 1998. The politics of technology: On bringing social theory into technological design. Science, technology & human values, 23(4), pp. 456-490.

Biesta, G. 2013. Responsive or Responsible? Democratic Education for the Global Networked Society. Policy Futures in Education, 11 (6), pp. 733–744.

Biesta, G. 2015. Good Education in an Age of Measurement: Ethics, Politics, Democracy. London: Routledge.

Biesta, G. 2016. Beyond Learning: Democratic Education for a Human Future. Taylor & Francis.

Biesta, G. 2019. What Kind of Society Does the School Need? Redefining the Democratic Work of Education in Impatient Times. Studies in the Philosophy of Education, 38 (6), pp.657-668.

Boellstorff, T. 2013. Making big data, in theory. First Monday, 18(10).

Borgman, C. 2015. Big Data, Little Data, No Data. Cambridge, MA: The MIT Press.

Bulger, M. Personalized learning: The conversations we're not having. Data and Society, 22(1) pp. 1-29.

Couldry, N. 2003. Digital divide or discursive design? On the emerging ethics of information space. Ethics and information technology, 5(2), pp.89-97.

Crawford, K., Miltner, K. and Gray, M.L. 2014. Critiquing Big Data: Politics, Ethics, Epistemology. International Journal of Communication 8, pp. 1663–72.

Crawford, K et al. 2019. AI Now 2019 Report. New York: AI Now Institute, 2019, https://ainowinstitute.org/AI_Now_2019_Report.html.

Davies, H.C., Eynon, R. and Salveson, C. 2020. The mobilisation of AI in education: a Bourdieusean field analysis. Sociology.

Davies, C. and Eynon, R. 2013. Teenagers and technology. London: Routledge.

Decuypere, M., 2019. Researching educational apps: ecologies, technologies, subjectivities and learning regimes. Learning, Media and Technology, 44(4), pp.414-429.

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

Driscoll, K. 2012. From Punched Cards to "Big Data": A Social History of Database Populism, communication +1: 1(4).

Espeland, W.N. and Stevens, M.L. 2008. A Sociology of Quantification. European Journal of Sociology. 49 (3), pp. 401-436.

Eynon, R. 2013. The rise of Big Data: What does it mean for education, technology, and media research? Learning, Media and Technology, 38(3), pp.237–240.

Fischer, C., Parados, Z. and Baker, R. 2020. Mining Big Data in Education: Affordances and Challenges Review of Educational Research 44(1), 130-160.

Gray, C.M., Boling, E. 2016. Inscribing ethics and values in designs for learning: a problematic. Educational Technology Research and Development 64, pp. 969–1001.

Hakimi L., Eynon R., and Murphy, V. (conditional accept) The ethics of using digital trace data in education: A thematic review of the research landscape. Review of Educational research.

Hartman-Caverly, S., 2019. Human nature is not a machine: On liberty, attention engineering, and learning analytics. Library Trends, 68(1), pp.24-53.

Jarke, J. and Breiter, A. 2019. Editorial: the datafication of education, Learning, Media and Technology, 44(1), pp. 1-6.

Kitchin R. 2014. Big Data, new epistemologies and paradigm shifts. Big Data & Society, 1(1)1-14.

Licenberg and Eynon (under review) Crafting Worldly Spaces: The Role of the Educator in Shaping Student Agency in Minecraft. Review of Education, Pedagogy, and Cultural Studies

Lipman, P., 2002. Making the global city, making inequality: The political economy and cultural politics of Chicago school policy. American Educational Research Journal, 39(2), pp.379-419.

Luckin, R., Holmes,W., Griffiths, M. and 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 & Technology, 44(1), pp. 77–86.

Macgilchrist, F. Allert, H. and Bruch, A. 2020. Students and society in the 2020s. Three future 'histories' of education and technology, Learning, Media and Technology,45:1, pp. 76-89,

Markham, A.M 2013. Undermining 'data': A critical examination of a core tem of scientific inquiry, First Monday, 18(10).

O'Neil, C., 2016. Weapons of math destruction: How big data increases inequality and threatens democracy. Broadway Books.

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

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

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

Powers, J.M., Fischman, G.E and Berliner, D.C. 2016. Making the Visible Invisible: Wilful Ignorance of Poverty and Social Inequalities in the Research-Policy Nexus. Review of Research in Education, (40)1, pp. 744-776.

Reay, D. 2011. Schooling for Democracy: A Common School and a Common University? A Response to “Schooling for Democracy”. Democracy and Education, 19 (1).

Regan, P.M. and Jesse, J., 2019. Ethical challenges of edtech, big data and personalized learning: twenty-first century student sorting and tracking. Ethics and Information Technology, 21(3), pp.167-179.

Roberts-Holmes, G. and Bradbury, A. 2016. The datafication of early years education and its impact upon pedagogy. Improving Schools.

Robins, K. and Webster, F. 1989. The technical fix: Education, computers, and industry. Macmillan.

Romero, C. and Ventura, S., 2020. Educational data mining and learning analytics: An updated survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3).

Saltman, K. J. 2016. Corporate Schooling Meets Corporate Media: Standards, Testing, and Technophilia. Review of Education, Pedagogy, and Cultural Studies 38 (2): 105–23.

Sefton-Green, J., 2020. Educating for Democracy in the Digital Age. In Oxford Research Encyclopedia of Education.

Selwyn, N., 2019. What’s the problem with learning analytics? Journal of Learning Analytics, 6(3), pp.11-19.

Spielman, A (2017) HMCI's commentary: October 2017. Available at, https://www.gov.uk/government/speeches/hmcis-commentary-october-2017.

Skinner, B.F., 1961. Teaching machines. Scientific American, 205(5), pp.90-106.

Wagner, B., 2018. Ethics as an escape from regulation: From ethics-washing to ethics-shopping. Being profiling. Cogitas ergo sum, pp.1-7.

Watters, A., 2014. The monsters of education technology. Ebook.

Webster, F., 2014. Theories of the information society. London: Routledge.

Williamson, B. 2017. Who owns educational theory? Big data, algorithms and the expert power of education data science. E-Learning and Digital Media, 14(3), pp. 105–122.

Williamson, B., 2021. Making markets through digital platforms: Pearson, edu-business, and the (e) valuation of higher education. Critical Studies in Education, 62(1), pp.50-66.

Winters, N., Eynon, R., Geniets, A., Robson, J. and Kahn, K., 2020. Can we avoid digital structural violence in future learning systems? Learning, Media and Technology, 45(1), pp.17-30.

Zeide, E., 2017. The structural consequences of big data-driven education. Big Data, 5(2), pp.164-172.

Zuboff, S. 2019. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. Profile Books: London.