PROJECT TITLE: Data-Smart schools - enhancing the use of digital data within secondary schools.

AIMS AND BACKGROUND

Aims

Contemporary Australian schools find themselves propelled 'into an era of dashboards and profiles, management and modelling, analytics and prediction' (Selwyn *et al.* 2018, p.161). While all elements of school communities are now expected to be making data-informed decisions, there is a current tension between that and the fact that much of the available, highly useful data remains inaccessible, underutilised or not used at all. It is a deceptively straightforward problem – while large quantities of digital data are now being generated within schools, much of this data remains poorly used and understood by school managers, teachers, students and parents.

The overarching aim of this project is to develop a realistic understanding of how digital data can be better used within schools to support decision-making and improve outcomes. Acquiring this knowledge is essential to better address the challenges of ensuring the continued effectiveness of Australian schools in the 'digital age'. Uniquely, the project aims to do this for whole-school communities – i.e. school managers, teachers, students and parents. The need for this project arises from the growing significance of digital data (generated both by institutions *and* by individuals) in shaping the 'everyday life of schools and classrooms' (Buckingham 2017, n.p).

This project will, through a combination of data-mapping, ethnographic case study and participatory design research, achieve the following objectives:

1. Map the full range of digital data being generated within schools related to educational processes – this includes data generated by the use of institutional management systems, online communications and IT infrastructures, as well as data generated by classroom apps, personal devices and other individual technology uses.
2. Work with school managers, teachers, students and parents to develop innovative digital processes, practices and tools that make more effective and empowering uses of existing digital data.

The project therefore expects to generate rich insights into the technical, informatic, organisational and social issues surrounding the (re)use of digital data in schools - engaging school managers, teachers, students and parents in developing models of digital data 'best practice' that address educational issues and problems. Given the importance now being attached to this topic elsewhere in the world, developing such understandings is essential to addressing the challenges of ensuring the continued effectiveness of Australian schools in the 'digital age'.

Background

Schools are now replete with digital devices, applications and systems, meaning that large quantities of data are generated, collated and processed through the use of these technologies on a daily basis. These data relate to most elements of schooling - ranging from the individual actions of students and teachers to whole school processes and 'performance'. While some of this data is 'officially' collected by school leaders and managers with analytic purposes in mind, vast amounts of 'naturally occurring' data are also generated through the daily use of school systems, personal devices and other technologies found in schools. Although schooling has always involved the collection of measurements, observations and statistics, these current forms of digital data are proving distinct in terms of the exhaustive amounts of detailed data that are now generated, the speed and flexibility in how this data is produced and processed and the range of data types and sources that now exist (Kitchen 2014).

This project is designed to address a growing tension within contemporary schools. On one hand, while the data being generated through the use of digital technologies is clearly of potential benefit to all constituent elements of a school, much of it remains underused or even not used at all. On the other hand,

all elements of school communities are now expected to be making use of data. For example, there is increasing expectation for school managers and administrators to support the more 'effective' running of schools through data-driven performance management and decision-making (see Decuypere *et al*., 2014). Similarly, teachers are expected to make use of various forms of learning analytics and student indicators to guide classroom activities, while students are encouraged to make decisions through self-tracking and self-quantification. In addition, parents are now expected to engage with school-produced performance data arising from their children's schooling.

Despite such imperatives, concerns are growing over how digital data are *actually* understood and used within schools. The existing academic literature addressing the realities of school digital data use has been led by a burgeoning community of scholars in Europe and the US, alongside our own research in Australia. In brief, this research generally has found in-school uses of digital data to be implicitly regulatory in nature and/or restricted in accessibility (e.g. Decuypere *et al.* 2014, Roberts-Holmes 2015). Thus, it has been observed that most digital data germane to improving educational processes and outcomes remain inaccessible to the majority of people within any school community. This exclusion is particularly pronounced for those outside of school administrative structures, such as classroom teachers, students and parents. The need for expanded data access and use within schools is therefore now beginning to be acknowledged (Graham *et al.* 2014).

As might be expected, the existing academic literature also points to a range of organisational, technical and social impediments to enhancing in-school uses of digital data. For example, previous studies have highlighted the limited quality, scope and inter-operability of digital data being generated in schools – especially with regards to the practical re-use of this data for other purposes (Selwyn *et al*. 2017). Most schools have limited capacity to develop their use of digital data, often relying on individual members of staff to act in *ad hoc* 'data manager' roles concerned largely with external data (such as NAPLAN scores) rather than the 'local' data being generated within each school (Breiter 2016, Selwyn *et al.* 2016). In this sense, school staff and students are found to often lack practical data skills and critical 'data imaginations' of how the data being generated through their engagement with digital technologies might be used differently (Finn 2016, Selwyn *et al.* 2017). It is noted that many schools are also reliant on digital systems and software configured for business rather than educational contexts, and often designed to support the transfer of school data to third parties rather than for in-school purposes (Williamson 2017). School leaders also face uncertainty over issues such as data privacy, data ownership and ethics, prompting many to take a cautious approach to their use of data (Lindh & Nolin 2016, Watson and Christensen 2017).

This emerging research base therefore provides the impetus for our DP19 project. We intend to investigate how digital data is being generated, accumulated and used within three diverse school settings. These findings will then be used to work with groups of school managers, teachers, students and parents in designing new processes, practices and tools that make more effective and empowering uses of the digital data being amassed in their schools. While mindful of the organisational, institutional and technical complexities of this task, it is believed that many of the issues currently restricting the potential of digital data can be addressed to allow for wider and more effective uses. This project is designed to discover the extent to which this is achievable.

PROPOSED PROJECT QUALITY AND INNOVATION

Ensuring the better use of digital data by 'ordinary' users is increasingly recognised as a significant issue across all areas of society. Popular and political understandings have moved from initial exaggerated hopes and fears in the early 2010s over the idea of 'big data', to more grounded current debates over data transparency, data rights and the broader accountability of data systems to the publics that they serve. Well-publicised incidents in areas such as advertising and policing have prompted concerns over the 'hidden' role that automated data-systems and algorithmic processes have in individuals' lives (O'Neil 2017) – specifically how lack of engagement with personal digital data impacts disproportionately on

already disadvantaged and excluded groups (Eubanks 2018). In short, enhancing public understanding of, and widening engagement with, digital data is beginning to be recognised as a significant problem across society. In this sense, our project will make a valuable contribution exploring how these broader concerns relate specifically to educational contexts.

Our project draws on two complementary theoretical approaches. First is the emerging social science field of 'critical data studies' (CDS) that strives toward the equitable and democratic engagement with digital data (Dalton *et al.* 2016). As such, our project is grounded in the CDS emphasis on advancing understandings of the social construction of digital data – i.e. the idea that digital data is socially created and has "a social life" (Lupton 2015, p.93). Our project therefore responds to the growing emphasis being placed within CDS on researching alternative ways "in which power and participation [might be] constructed and enacted" in individuals' data practices (Couldry & Powell 2014, p.1). This has already seen CDS studies (including work by Selwyn & Pangrazio) that investigate how data generated from personal uses of digital technologies might be repurposed to enhance individuals' data agency – i.e. 'feeding such data back to users, enabling them to orient themselves in the world' (Kennedy & Moss 2015, p.1).

Secondly – and complementing our theoretical interest in CDS – is the fast-evolving area of learning analytics. Learning analytics (LA) and educational data-mining (EDM) has developed over the past decade, bringing computational science and data science techniques to bear on the use of data in educational contexts (Siemens 2013). In particular, this field is at the cutting-edge of advancing the use of digital data relating to students, teachers and their immediate educational contexts to enhance learning, pedagogy and organisational processes (Wilson *et al.* 2017). Specific areas of attention include the use of digital data to inform individual and institutional sense-making, decision-making and subsequent behaviour change (Ferguson 2012). Learning analytics therefore makes use of techniques from the computational and data sciences – from web analytics and business intelligence to data-mining, machine learning and artificial intelligence. Researchers (including ground-breaking work by CI Gašević) have worked to make optimal use of digital data arising from learning management systems, management information systems, student feedback, attendance data and other sources of data generation. This includes developing tools and techniques relating to data visualisation, data integration, predictive modelling, and classroom orchestration. This field of work therefore frames the use of digital data in schools in terms of technical, informatic and social factors, and will play a key role in guiding our project to translate the social and political concerns raised by CDS into practical data-based changes in schools.

Research questions

Against this background, the project will address the following four research questions:

RQ#1 - What digital data is being generated through the use of digital technologies within schools? For example, what data is being generated by school systems, as well as individual staff and student use of technology? What aspects of school education does this data relate to? In what forms does this data exist and in what forms are these data accessible (e.g. open/closed access, raw/value-added)? Is this data created intentionally or 'naturally-occurring'? What is the quality, scope, inter-operability and compatibility of these data?

RQ#2 - What are the current uses of this digital data within school communities? How is this digital data being used and by whom? For instance, how is this data being used to inform decision-making and planning (both for individuals and the institution)? What is the 'social life' of this data – i.e. how is digital data being re-used beyond its primary intended purposes? Is this data work conducted primarily by administrative elites, or individual teachers, students and parents? What evidence is there for innovative data practices? What data skills, competencies and interests exist across staff, student and parent groups? What digital data is *not* being used, and why? What technical, organisational and social issues are shaping data use within schools?

RQ#3 - How might this digital data be more effectively accessed and used in schools? What new uses of the digital data currently being generated in schools might be of benefit to school managers, teachers, students or parents? What institutional processes, individual practices and software-based tools are necessary to achieve these aims? For example, how can this digital data be made widely accessible to staff, students and parents? What policies and processes can be developed to increase the quality of datasets (in terms of scope, interoperability and compatibility)? What individual skills and understandings need to be developed for people to engage meaningfully with the school-related data that is being generated through digital technology use?

RQ#4 - What are the consequences of 'enhanced' uses of digital data within schools? What are the outcomes when changes are made to the ways in which this digital data is used within schools? Which new data processes, practices or tools lead to improved outcomes within school? Do different ways of using data alter social relations within schools? Does different ways of using data lead to new relationships and understandings? Does different ways of using data lead to better decision-making amongst staff, students or parents? What technical, organisational and social issues continue to shape the (non)use of digital data within schools?

Research design and methods

The scope of these questions calls for in-depth, qualitative research that is sensitive to different school contexts. As such, this project takes the form of a comparative case study of three contrasting government secondary schools in Victoria. Within each of these schools, the research questions and underpinning aims of the project will be addressed through three phases of iterative research activity:

Phase 1 – Mapping the 'data landscapes' and existing uses of digital data in the three case study schools [Months 7-12, RQ1 & RQ2]

i) Data audit within each school

Phase 1 of the project will first involve the research team conducting thorough 'data audits' of each school. The audits will identify the nature and form of digital data sources and the conditions of access. These are likely to include data generated from learning and assessment management systems, school administration systems, email and other communication systems, server data and other IT infrastructure sources, classroom uses of technology, personal devices and so on. It is anticipated that this initial stage of the research will involve in-depth auditing (i.e. identification, assessment and cataloguing) of data sources, alongside 16 in-depth interviews with key actors in each school (leadership teams, IT departments, key administrators and leading teachers) and 10 interviews with external data providers (platform providers, external IT contractors, IT and data teams from Department of Education). These activities will be based around the ODRN (Open Data Research Network) framework for 'assessing data supply', and will provide a comprehensive map of the quality, scope and interoperability of these digital data sources within each of the schools.

ii) Ethnographic case studies of current digital data use within each school

These 'data audits' will then frame the in-depth study of the current uses of digital data within each of the three schools. Here the research team will conduct sustained ethnographic case studies of data (non)use within each school over Terms 3 and 4 of the 2019 academic year (one day per week per school site). In order to ensure the timely and successful completion of this phase, research activities within each school will focus on data associated with three relatively robust digital data sources identified from the data audit (for example, data generated from the school LMS, email systems, IT infrastructure use, popular data-gathering tools such as Google Forms). Using an organisational ethnography approach (see Ybema *et al.* 2009), ethnographic techniques such as interviews, observations, extended field notes, and document and policy analyses will be employed to gain a

detailed, 'thick' sense of how different actors encounter and experience this digital data – paying specific attention to the institutional processes, individual practices and software-based tools relating to data use in each school. Alongside face-to-face methods, 'trace ethnography' techniques will be used (Geiger & Ribes 2011) to follow examples of these three different types of data across school and staff/student/parent digital spaces. The three types of data will be selected to be of relevance either to school managers (e.g. related to resource use/allocation); teacher and students (e.g. related to classroom activities) and parents (e.g. related to home/school communications). These data will be analysed using established methods from the computational social sciences such as data mining, usage data interrogation techniques, organisational network and social network analyses.

Phase 2 – Participatory design: designing new data processes, practices and tools [Months 13-24, RQ3]

The second phase of the study will then explore questions of how digital data use might be enhanced in each of the three schools. Here the project will use a 'critical participatory design' approach which will involve small groups of staff, students and parents in the conceptual design of data processes, practices and tools that better reflect their interests and needs (Bodker 1993). In order to ensure the timely and successful completion of this phase of the project, research activities within each school will continue to focus on the three digital data sources identified in each school from Phase 1. Sampling of individuals for these participatory design groups will aim to recruit groups who are broadly representative of staff, teacher and parent populations in each school.

In each school, a series of participatory design workshops will be run with groups of students, managers, teachers and parents (6 to 8 participants in each group). The aim of each group will be to cooperatively design data processes, practices and tools that meet self-identified needs. From previous research, it is anticipated that these different groups will have diverse uses for the digital data – ranging from decisions relating to teaching and learning, through to time management and other logistical concerns. One of the initial aims of the workshops, therefore, is to explore what areas of 'school' that different individuals and groups are interested in using data for, and what outcomes are realistically achievable. In this sense, the workshops will focus on developing data processes, practices and tools that are: (a) related to educational activities, and (b) can be deemed to involve what Ruppert (2016) terms 'decisive data' – i.e. digital data that is mobilising in its effects and supports actual change and tangible outcomes that are meaningful to the individuals concerned.

The series of ten two-hour workshops will therefore cover the following sequence of sessions:

- Four x 'Data Literacy' workshops: initial sessions led by the research team to increase the data awareness, and skills of participants, as well as establish working relationships between all members;
- Two x initial 'Exploration and Co-interpretation' workshops: gaining familiarity with the available data-sets; discussing and identifying 'real world' school-related problems that the available data-sets might relate to;
- Two x 'Discovery Processes' workshops: co-designing data processes, practices and tools that respond to these interpretations;
- Two x 'Prototyping' workshops: developing and evaluating corresponding designs and solutions for new data processes, practices and tools.

These workshops will follow the participatory design progression from 'initial exploration and co-interpretation' sessions, through to 'discovery processes' and then 'cooperative prototyping' sessions (Spinuzzi 2005). Over the ten weeks, the workshops will be encouraged to focus on three different forms of 'designs and solutions', i.e.:

- Data Processes: i.e. prototype designs for re-configuring technologies and systems that generate digital data; altered permissions to access data; revised school policies and guidelines relating to

data access and use;
- Data Practices: i.e. prototype designs for skills, understandings and techniques that individuals can draw upon when engaging with the digital data;
- Data Tools: i.e. prototype designs for bespoke 'simple' apps or web-based tools that allow participants to view and/or work with digital data sources.

Using protocols developed in our previous research, these workshops will be appropriate for participants with no prior computational experience, following the approach of the 'data carpentry' tradition (Teal *et al.* 2015). As such, the workshops will aim to develop participants' skills in retrieving, viewing, manipulating, analysing, visualising and sharing data in order to understand the value of this data to their school-related activities.

These workshops will generate empirical considerable data. The researchers will systematically collect empirical data from the workshops in the form of observations, recorded discussions and analyses of the artefacts from the design processes (e.g. sketches, plans, storyboards). The research team will act as scribes for the drawing-up of the final designs. Each workshop group will work towards the goal of producing prototype designs for one 'data process', one 'data practice' and one 'data tool'. A programmer will also be employed for this component of the project to advise the groups and to begin developing the data 'tool' prototypes for use in Phase 3.

Phase 3 – Field-testing new data processes, practices and tools [Months 25-34 – RQ4]

The final phase of the investigation will involve testing how the proposed data processes, data practices and data tools work when implemented in the realities of the school context. In each of the three schools, the research team will work with school leadership and IT teams to see which of the proposed 'data processes' might be enacted (i.e. new policies, system configurations, access permissions). The research team will also work with workshop participants to publicise and support the adoption of the proposed 'data practices' throughout the wider staff, student and parent populations. Finally, the research team will finalise the building and beta-testing of one of the proposed 'data tools' from each of the three schools (likely to take the form of mobile-based apps, web-based programs or equivalent). These will be three tools that are considered to be the most technically feasible in terms of software development. A programmer will be employed for this component of the project due to the complexity of aggregating data from different systems into usable and accessible tools.

These new data processes, practices and tools will then be implemented and evaluated in each of the respective schools during the remainder of the school semester, thereby gaining an initial sense of how these new ways of engaging with digital data 'work' *in situ*. Each of the three new data tools will be instrumented, allowing the collection of additional trace data to track their usage across the schools. Researchers will make ten follow-up visits across the semester to each case study school – focusing initially on managers, teachers, students and parents from the participatory workshops, and then other user groups of the data processes/practices/tools as they emerge. Interview and observation data collected during these visits will help develop a sense of how different groups within each school engage with the data processes, practices and tools *in situ*, and what organisational, technical and social factors mitigate the success (or otherwise) of these changes.

FEASIBILITY

This project is innovative and ambitious, and arises from a number of projects successfully completed by the investigators that give us confidence in the research design, methods and approach. On the basis of our experience from these previous studies, the proposed project has been planned carefully to be conducted and completed within a realistic 36 month timeframe and an efficiently costed budget. The

three-stage iterative mixed-methods design of the project has been developed to ensure the production of a large volume of high-quality empirical data. The conduct of the project will adhere to the following timeline (starting Jan 2019):

- Months 1 to 6 - Preparation: planning, development of research instruments, first visits to study schools
- Months 7 to 12 - Phase 1(1): 'Data audits' in three case study schools
- Months 13 to 18 - Phase 1(2): Examining current digital data practices
- Months 19 to 24 - Phase 2: Participatory Design: designing new data processes, practices and tools
- Months 25 to 30 - Phase 3: Testing new data processes, practices and tools
- Months 31 to 36 - Final empirical data analysis, writing-up and dissemination activities

As such, the project has been designed to be feasible in terms of the team, its design, budget and timeline. Through our research projects preceding this DP19, the team has developed a high level of expertise both in the empirical data collection methods to be used and the development of digital data tools and techniques. In terms of research design, the project has been designed with clearly manageable boundaries – not least the focus of latter phases on three specific data sources in each school. Similarly, given the time limits of the project, it is expected the three digital tools that will be developed to be very simple in nature - aiming primarily to achieve 'proof of concept' rather than developing new fully-functional applications. Our project is therefore intended to provide valuable initial 'prototype' information to underpin later larger-scale changes.

This research will be located within a high quality and supportive research environment within Monash University's Faculty of Education and its newly established Education Futures research centre. As such the project will benefit from the following sources of support:

- **Faculty:** The Faculty of Education has a world-leading reputation for school-based research and an emphasis on developing and maintaining a critical mass of leading professors and researchers. The 2017 QS World University Rankings ranked Monash University 17th in the world for the study area of Education.
- **Faculty Research Centre:** This project will be based in the Faculty's newly established Education Futures research centre. This is an institutionally-funded centre supporting four strands of activity, including work on Digital Futures (led by CI Selwyn) and Learning Analytics (led by CI Gašević). This DP19 would be one of the first government-funded projects to be based in the Education Futures centre, and would have full benefit of the centre's administration and project management facilities.
- **Faculty research group:** The Digital Education Research group at Monash (previously Learning with New Media) has been at the forefront of social research into new media use in education for over 20 years, and has been throughout its life a highly collegial and collaborative group that has developed many leading educational technology scholars now in other Australian and international institutions.

BENEFIT

Enhancing digital data understandings and use amongst non-expert users is a fast-growing area of research, and this would be the first large-scale empirical research project to address this issue with regards to schools and education. The project applicants have established a well-recognised track-record in the different areas that the project aims to address, and this project will further advance their contribution to the knowledge base. The likely benefits of this knowledge are considerable. Enhancing digital data skills and understandings amongst school workforces and students will both be advantageous to the Australian economy and improve the institutional effectiveness of schools (McKinsey 2013). As such, the project is designed to be of considerable real-world significance to secondary schools and the

wider educational sector, while costed to make cost-efficient use of existing technological resources and open data principles. It is intended that the main outcome of the project will be the enhanced understanding of the nature of school practice and improvement in the context of rapidly burgeoning policy concerns over big data, data mining and the increased use of digital technologies. These are educational issues that are becoming recognised as important around the world. This project will position Australian educational research at the forefront of these debates.

The project is designed to generate meaningful findings for a range of end-users beyond the three study schools. Given the burgeoning interest in data studies, it is expected international academic take-up of the project outputs to be high. In particular, the project's findings and outcomes will therefore be of international policy interest – addressing directly current policy concerns in Europe, North America and east Asia. Alongside ongoing development of commercial and open 'learning analytics' systems, the notion of 'data-driven decision making' is now a common imperative in East Asian, North American and European schools. This trend is being replicated in Australia. For example, the Victorian Government's 'Towards Victoria as a Learning Community' committed to making it 'easier for data to be shared with students, parents and teachers'. This research project offers a school-focused perspective on the issues underlying this under-use of digital data, and a set of practical, cost-effective but powerful school-focused solutions.

It is also intended that the project will be of clear significance to practitioners and other organisations that work in the area of schools and schooling. In terms of value for money, the source code of the data tools produced for Phase 3 will be published under an open source license (e.g. Apache) in an open repository (e.g. GitHub). The project has been designed to provide a range of practically useful insights for school managers and managers, advocacy groups and non-governmental public organisations, organisations and professionals that work with educational technology, and IT industry actors that play a key role in the resourcing and support of school technology systems. The project aims to demonstrate, for example, how digital data can best fit within school settings, fitting with a number of concurrent concerns (e.g. with school improvement, performance indictors and measurement, education analytics). In particular, in terms of its significance to education professionals, and those that work with educational professions the project has potential implications for educational practice, where issues relating to teaching training, pedagogy, school leadership and management are explicitly addressed. As such, the research will provide valuable information to professional groups as well as general teaching unions and school leadership associations.

In terms of academic benefit, the outcomes of this project will be of value to researchers in several fields:

- For those in the field of Educational Research, the project will develop the analysis of school change and innovation via a more elaborate account of the role of digital technology in school and its role in generating data relating to educational processes. The project will also help to feed into broader understandings of the changing nature of educational institutions and the professional practices of teachers and the learning practices of students in the 'data age'.
- For those in the field of Learning Analytics and EDM, the research will provide much needed research on the applicability of LA strategies in K-12 schools, and with school managers and parents (to date under-represented in the LA literature). In addition, the project's focus on the (re)use of already existing school data will offer a valuable counterpoint to the traditional focus within LA on development of specialised analytic and adaptive learning systems.
- In terms of Critical Data Studies, the project will be one of the first empirical education research contributions to this emerging academic field – extending work from other areas on the socially-shaped nature of data work in institutional settings.

REFERENCES

Bødker, S. (1993). Cooperative design. In Schuler, D. (Eds.), *Participatory design*, Hillsdale, NJ: Lawrence Erlbaum (pp.157-175)

Breiter, A. (2016). Datafication in Education. In *Int. Conf on Stakeholders & IT in Education* (pp. 95-103). Springer, Cham.

Buckingham, D. (2017) *Learning in the matrix: the datafication of education*. www.davidbuckingham.net., December 18th

Couldry, N. and Powell, A. (2014). Big data from the bottom up. *Big Data & Society*, *1*(2)

Dalton, C., Taylor, L. and Thatcher, J. (2016) Critical data studies. *Big Data & Society*, *3*(1)

Decuypere, M. Ceulemans, C. and Simons, M. (2014) Schools in the making. *Journal of Education Policy*, 29(5):617-639

Eubanks, V. (2018). *Automating inequality*. New York, Macmillan

Ferguson, R. (2012). Learning analytics. *International Journal of Technology Enhanced Learning*, *4*(5-6):304-317.

Finn, M. (2016). Atmospheres of progress in a data-based school. *Cultural Geographies*, *23*(1):29-49.

Geiger, R. & Ribes, D. (2011). Trace ethnography. in *System Sciences (HICSS), 44th Hawaii International Conference on* (pp. 1-10). IEEE.

Graham, C. *et al.* (2014). Data-driven decision making in the K-12 classrooms. in Spector, J. (ed) *Handbook of Research on Ed Communications and Technology*. New York: Springer (pp. 337-349)

Kennedy, H., & Moss, G. (2015). Known or knowing publics? *Big Data & Society*, *2*(2),

Kitchin, R. (2014). *The data revolution.* London, Sage

Lindh, M. & Nolin, J. (2016). Information we collect. *European Educational Research Journal*, *15*(6), 644-663.

Lupton, D. (2015). *Digital sociology*. London, Routledge

McKinsey (2013). *Open data*. McKinsey Global Institute Report.

O'Neil, C. (2017). *Weapons of math destruction*. New York, Broadway Books.

Roberts-Holmes, G. (2015). The 'datafication'of early years pedagogy. *Journal of Education Policy*, *30*(3):302-315.

Ruppert, E. (2016). Big Data Economies and Ecologies. In: L. Ryan and L. McKie, eds. *An End to the Crisis of Empirical Sociology?*. London: Routledge (pp. 12-26).

Selwyn, N., Henderson, M., & Chao, S. (2015). Exploring the role of digital data in contemporary schools and schooling. *British Educational Research Journal*, *41*(5):767-781.

Selwyn, N., Henderson, M., & Chao, S. (2017). The possibilities and limitations of applying 'open data' principles in schools. *Cambridge Journal of Education*, *47*(2):167-187.

Selwyn, N., Nemorin, S., Bulfin, S., & Johnson, N. (2018). *Everyday Schooling in the Digital Age*. London, Routledge.

Siemens, G. (2013). Learning analytics. *American Behavioral Scientist*, *57*(10):1380-1400.

Spinuzzi, C. (2005). The methodology of participatory design. *Technical Communication*, 52(2):163-174

Teal, T. *et al*. (2015). Data carpentry. *International Journal of Digital Curation*, 10(1):135-143.

Watson, R. & Christensen, J. (2017). Big data and student engagement among vulnerable youth. *Current Opinion in Behavioral Sciences*, *18*:23-27.

Williamson, B. (2017). *Big Data in Education*. Sage.

Wilson, A. *et al.* (2017). Learning analytics: challenges and limitations. *Teaching in Higher Education*, 22(8): 991-1000

Ybema, S. *et al*. (2009). *Organizational ethnography*. London: Sage.