Critical literacies for a datafied society: academic development and curriculum design in higher education

Javiera Atenas\textsuperscript{a,b,*}, Leo Havemann\textsuperscript{c,d}, and Cristian Timmermann\textsuperscript{e}

\textsuperscript{a}Faculty of Education, Universitat de Barcelona, Barcelona, Spain; \textsuperscript{b}Latin American Initiative for Open Data, London, UK; \textsuperscript{c}Institute of Educational Technology, The Open University, Milton Keynes, UK; \textsuperscript{d}Digital Education, University College London, London, UK; \textsuperscript{e}Interdisciplinary Centre for Studies in Bioethics, University of Chile, Santiago, Chile

Received: 27 May 2020; Revised: 3 August 2020; Accepted: 19 August 2020; Published: 05 October 2020

Participation in democracy, in today’s digital and datafied society, requires the development of a series of transversal skills, which should be fostered in higher education (HE) through critically oriented pedagogies that interweave technical data skills and practices together with information and media literacies. If students are to navigate the turbulent waters of data and algorithms, then data literacies must be featured in academic development programmes, thereby enabling HE to lead in the development of approaches to understanding and analysing data, in order to foster reflection on how data are constructed and operationalised across societies, and provide opportunities to learn from the analysis of data from a range of sources. The key strategy proposed is to adopt the use of open data as open educational resources in the context of problem- and research-based learning activities. This paper introduces a conceptual analysis including an integrative overview of relevant literature, to provide a landscape perspective to support the development of academic training and curriculum design programmes in HE to contribute to civic participation and to the promotion of social justice.

Keywords: critical pedagogy; critical data studies; open education; open data; open government; social justice; civic education; human rights; politics; government; policy; capacity building

Introduction

We live in a ‘datafied’ society where almost everything is continuously transcribed into data, quantified and analysed (Van Es and Schäfer 2017), where decisions taken by corporations and governments are increasingly data- and algorithm-driven. Such data are often said to be ‘collected’ (as if pre-existing, and therefore, simply reflecting reality), but the processes through which data are generated, communicated and represented are neither necessarily transparent nor devoid of negative effects. The promised utopia of Big Data has been revealed as an actual dystopia of Big Brother: data-based ‘surveillance capitalism’ has increasingly been adopted as the leading business model of our information age (Zuboff 2015), and the simultaneous ubiquity and
opacity of corporate data practices often have a deliberately ‘chilling’ effect on public critical engagement (Draper and Turow 2019).

According to Dalton, Taylor, and Thatcher (2016), people are often unaware why, how or even that data about themselves are being collected, analysed and ‘shared’ with additional parties. For Ozga (2008) and Ball (2015), the data we produce are used to classify us as individuals, allocating us into categories that define our worth in society, our perceived effectiveness and our potential to pose risks. Using algorithms, our behaviours, successes and failures are then predicted according to the categories in which we have been placed (Harel Ben Shahar 2017; Maull, Godsiff, and Mulligan 2014; Schildkamp, Karbautzki, and Vanhoof 2013; Schouten 2017).

In parallel with monitoring and analytics undertaken by external agencies, people are also now increasingly engaging in ‘voluntary’ digitised self-tracking. This set of practices is currently at the centre of consumer interest and industry attention (Haddadi and Brown 2014) and often referred to under the umbrella of the quantified self. For Lupton (2016), such ‘opt-in’ forms of ‘dataveillance’ are ‘now being used in situations where the choice to participate may be limited’ as ‘self-tracking practices are advocated and implemented in many social contexts and institutions’ (p. 103). Beer (2018) notes that this quantification of labour has been associated with and enabled a new precarity in working conditions. According to Moore (2017), such inequalities are worsened by a widespread lack of access to educational opportunities to understand and challenge datafication, as even higher education (HE) has been relatively uncritical of these phenomena of ruthless quantification.

Fostering critical understanding of data in HE is key, as education has been permeated and transformed by data and data practices, a process that has been described as ‘datafication’ of education (Gilliard 2017; Gulson and Sellar 2019; Lupton and Williamson 2017; Selwyn 2015). In HE, a perceived need to quantify learning has justified an ever-increasing metricisation of student and staff activity, re-casting students, educators and learning activities into roles as data-producers. This datafication of learning is providing a steady diet for learning analytics, one that is associated with diacatholicons promising to improve and enhance education through personalisation and data-driven interventions (Baker and Inventado 2014; Schouten 2017), at the risk of jeopardising complex and critical student-centred pedagogies (Roberts-Holmes 2015) and potentially coercing digital participation in order to ensure data are gathered (Barassi 2019).

In reference to the collection and analysis of data by HE institutions, Shacklock (2016) worryingly argues:

Growing up in a digital world dominated by Google, Facebook, and Amazon, the current generation of young undergraduate students are used to providing their personal data in return for access to services and products, and perhaps see their relationship with their university in the same light. (p. 38)

This statement is something of a red flag, as being ‘used to’ something neither equate to an informed understanding of data and privacy issues (Draper and Turow 2019; Hargittai and Marwick 2016) nor lend such practices legitimacy.

The case of students in HE illustrates a wider point that the acquisition of critical data literacies has a strong relevance for social justice; without them, opportunities to challenge dominant narratives and to influence and change social circumstances and practices will be curtailed. A data literacy divide will likely widen the gap between
privileged and underprivileged groups. While certain groups will be well positioned to participate economically and socially, even becoming decision-makers, those who cannot engage with data may remain or become further marginalised, ultimately only playing the role of data points, to be studied ‘from above’ (Atenas and Havemann 2019; Johnson 2014).

Hood and Margetts (2007) argue that governments operate through two sets of agents: detectors and effectors. Detectors gather information (data) from individuals and society, and effectors seek to influence them. The media and other powerful social actors such as corporations must also be understood to play both detecting and effecting roles in society. In order to understand how various organisations track and attempt to manipulate our habits, conduct, political views and relationships, citizens need to be trained to become detectors and effectors themselves. Educational and citizenship development programmes need to bridge civil society, industry, research and politics, promoting the effective and efficient use of data and information to critically participate in democratic and social dynamics (Rychen and Salganik 2003).

Van Es and Schäfer (2017) note that ‘students need to be educated to become critical data practitioners who are both capable of working with data and of critically questioning the big myths that frame the datafied society’ (p. 12). Therefore, academic capacity building programmes must look beyond data capabilities, and include critical thinking, citizenship and innovation skills (Gray, Gerlitz, and Bounegru 2018; Van Es and Schäfer 2017), foster skills to evaluate, analyse and interpret data (Prado and Marzal 2013; Schield 2004), and ground teaching and learning in addressing real problems (Atenas and Havemann 2019; Fung 2017), to understand issues such as commodification, surveillance and privacy (Gray 2016; Kellner and Share 2009; Matthews 2016). In light of this, it is crucial that academics and students across disciplines in HE adopt as foundational a critical data studies perspective (Iliadis and Russo 2016). Such an approach can empower students to question the ethics, structures and economics of data use, and fundamentally, the apparent inevitability of the surveillance and datafication of all aspects of daily life.

In order for students to appreciate and critically analyse ‘the societal embeddedness and constructedness of data’ (Richterich 2018, p. 2), we suggest that it is necessary to work with real data. This presents a challenge, as vast swathes of data remain off-limits in terms of public access. However, recent years have seen an increased focus on and demand for availability and transparency of public and research data produced by international organisations, governments, civil society, scientists and research centres. Furthermore, we argue that it is vitally important that students gain literacies for accessing and working with datasets which are not simply openly available, but which make important truth claims about beings and activities in the world. Such datasets are increasingly released as open data, which represent a nexus of challenge, opportunity and responsibility in HE, as open data can be used as open educational resources (OER) to support the development of research data literacies, transparent and open scientific practices, as well as citizenship and critical thinking, regardless of the research field (Atenas and Havemann 2015).

The use of open data as OER is an emergent strand within a wider realm of open educational practices (OEP) (Atenas and Havemann 2015). OEP are a ‘broad descriptor of practices that include the creation, use, and reuse of OER as well as open pedagogies and open sharing of teaching practices’ (Cronin 2017, p. 15), as well as practices which ‘respect and empower learners as co-producers on their lifelong learning path’ (Andrade et al. 2011). Thus, OEP can be an enabler to foster the
development of data skills by promoting students working with real problems through a critical engagement with raw and open data (Atenas and Havemann 2015).

Open data can be analysed by students as citizen researchers, empowering them to frame, ask and investigate socially pertinent questions, and thereby better exercise their rights as citizens (Arzberger et al. 2004; Atenas, Havemann, and Priego 2015; Huijboom and van de Broek 2011; Molloy 2011). Open data can be used across and between disciplines, by promoting a critical engagement with the same raw data that researchers, governments, civil society, international organisations and policy-makers generate and use, supporting the development of information, statistical, scientific, media and political literacies, and critical thinking, collaborative and citizenship skills (Atenas, Havemann, and Priego 2015; Johnson 2014; Manca et al. 2017; Markham 2018).

It is also important to foster coherent, holistic institutional strategies and policies in HE to support educators in developing data literacies, as educators need a solid ground upon which to build innovative pedagogical practices (Atenas et al. 2019), which can often be described as OEP as these are catalysts for design and iteration of complex, collaborative, participatory and transparent teaching and learning processes (Atenas et al. 2019; Conole and Ehlers 2010; Cronin 2017; Havemann 2016, 2020; Manca et al. 2017).

**Method**

To briefly recap the problem as outlined in the introduction, in an increasingly digital and datafied society, data are often accorded the status of objective fact, despite its constructed, partial and biased nature. Algorithms drawing upon data are used to profile members of society and make crucial decisions which likely disproportionately impact those with less privilege and resources at their disposal. Students entering HE are not necessarily aware of, let alone equipped to research, understand and analyse the resulting issues. Therefore, educators in HE should be supported in embedding the development of transversal data literacies, including both data practices and criticality, into curricula.

While we contend that an approach grounded in the use of open data as OER (Atenas and Havemann 2015, 2019) remains a key strategy, here we are seeking to expand the discussion of how and what students can learn through such critical open data studies, and propose that this question can only be effectively approached from multiple angles of analysis. The purposes of this discussion are as follows: first, to raise awareness of the need for critical data literacies (Markham 2018); second, to highlight the significance and relevance of openly available, real data for educational purposes; and third, to suggest strategies for fostering data literacies and practices that can be adapted for local and disciplinary contexts.

As authors we began by drawing upon some of the diverse disciplinary contexts we are collectively versed in: information, digital and media literacies, digital and open education, open data and applied philosophy. In order to scope the vast size of the potentially relevant literature dataset, we conducted a search using Google Scholar, which has been found to index a wider number of sources than other scholarly databases across a range of disciplines (Martín-Martín, Orduna-Malea, and Delgado López-Cózar 2018). The search terms used and resulting number of records retrieved were ‘data literacy’ (9860), ‘open data’ (255 000), ‘media literacy’ (146 000), ‘critical data studies’ (1230), ‘datafication’ (8710), ‘digital citizenship’ (16 900) and ‘education data’ (150 000).
It became clear to us that an emerging topic, such as the focus of this study, does not lend itself to a ‘systematic’ review of existing literature, instead requiring a more ‘open’, exploratory and interdisciplinary approach. Indeed, there is not, as yet, an extensive body of literature reporting on the outcomes of such activities to be reviewed; rather, we draw on a range of fields for guidance and inspiration in order to consider relevant arguments and evidence available to support the proposed approach. Our review of relevant literature is therefore better understood as exploratory or integrative in nature (Souza, Silva, and Carvalho 2010). As Torraco (2016) notes, integrative reviews address ‘new or emerging topics that would benefit from a holistic conceptualization and synthesis’ (p. 410).

The need for critical data literacies

The key educational challenge presented by the datafied society is an urgent requirement for transversal data literacies amongst educators and learners (Mandinach and Gummer 2013; Prado and Marzal 2013; Schield 2004; Stephenson and Shifter Caravello 2007). Drawing upon Foucault (1980), Giroux (2010) calls for an education which develops and improves people’s abilities to recognise and challenge power dynamics, enabling them to become a committed and critical citizenry, which is able to expand and deepen their participation in the promise of substantive democracies. To become such active citizens, people must develop transversal skills (UNESCO 2016), which are a comprehensive set of abilities for living and working, such as critical thinking and information, data, media and political literacies (Atenas, Have mann, and Priego 2015), to facilitate a long-term democratic commitment (Dudley and Gitelson 2010).

The technical and intellectual tools required to discover, interrogate and interpret datasets are neither widely nor equitably distributed throughout society, despite these being key to enable ‘active citizen participation in public decision processes, for influencing public policies and for social action’ (Engel 2017, p. 44). Failure to learn how to understand, analyse and challenge, data will result in citizens being in a continuously increasing position of informational disadvantage in relation to socio-political and commercial actors. Consequently, data literacy education needs to address a broad vision of data as social as well as technical assemblages, which include all of the ‘technological, political, social and economic apparatuses and elements’ that constitute and frame its production and use (Kitchin and Lauriault 2014, p. 6).

HE has traditionally played a key role in fostering data literacies and practices, but the current moment appears to require a renewed focus on these, which we suggest should be critically oriented and across disciplines. Meanwhile, the open data movement has argued for public data, in particular, to be made accessible and open for public interrogation to ensure transparency and accountability, but as Johnson (2014) argues, the fact that such datasets are openly licensed does not ensure that they come without the problems of reproduction of socially embedded biases; thus, it is a key to foster data capabilities amidst groups and individuals who might wish or need to analyse it, which clearly includes educators and students in HE.

This points to a need for curriculum design, which expressly links the interrogation of data with the possibility of socio-political action. Critical, socially engaged data literacy education can support the realisation of three ideals of social justice: (1) the right to self-determination, which requires the ability to make an informed decision, (2) the right to participate in science and cultural life, which in turn needs access to
data for its full realisation and (3) contributive justice, to facilitate an environment that encourages the use and development of individual and community capabilities (Timmermann 2014, 2017).

Academic development on critical data pedagogies

Fostering data literacies within HE requires interdisciplinary pedagogical efforts to design learning activities using problem- and research-based learning approaches (Atenas, Havemann, and Priego 2015; Maybee and Zilinski 2015; Raffaghelli 2018). For Mandinach and Gummer (2013), educators’ data literacy is ‘the ability to understand and use data effectively to inform decisions … composed of a specific skill set and knowledge base that enables educators to transform data into information and ultimately into actionable knowledge’ (p. 30).

Ebbeler et al. (2016) suggest that academic development must be grounded on collaborative learning and in solving problems, which should emerge from the educators’ own interests but that requires the support and active involvement of colleagues. Wayman and Jimerson (2014) argue that training for academics in data literacies needs to be contextual, coherent, resourced and sustainable, while Coburn and Turner (2016) emphasise the importance of providing expert guidance to help educators to develop data-led activities. Mandinach and Gummer (2016) suggest a data-driven workflow with five components: (1) identify problems and frame questions through data; (2) retrieve data; (3) transform data into information; (4) transform information into decision; (5) evaluate outcomes.

Educators’ data literacies can be developed to inform professional practices, including teaching and research evaluation (Rodés, Gewerc-Barujel, and Llamas-Nistal 2019). For Dunlap and Piro (2016), data literate educators can design effective data-based learning activities. According to Kippers et al. (2018), these educators are well placed to help learners to consciously collect and analyse data to construct knowledge; therefore, training programmes should foster research on pedagogical practices towards implementing data literacy activities whose impact can be assessed (Izadinia 2014; Phuong, Cole, and Zarestky 2017).

Halverson et al. (2007) propose to train educators in data projects with dynamics they can extrapolate to their courses, connecting data to support evidence-based approaches for teaching and learning, including activities such as data challenges, considering the guidelines of Piety, Court, and Hickey (2014) who suggest that data-led activities should include elements of ethics, privacy and information architecture.

Innovation in curriculum design: open data as OER

Open pedagogies, media education and data literacy, need to enable students to become data intermediaries (Davies and Edwards 2012; Magalhaes, Roseira, and Strover 2013), thereby making a positive impact on democracy. The adoption of open data as OER (Atenas and Havemann 2015, 2019; Coughlan 2019) could form a strong basis to raise awareness of socio-political issues, promoting learning through critical thinking, which Brookfield (1987) describes in three phases: (1) determine the hypothesis that guides decisions and actions; (2) verify the accuracy of these assumptions by analysing perspectives, data and information; (3) make informed decisions based on evidence and research.
Understanding data requires the development of a range of skills, and curricular design for fostering data literacies must therefore be carefully planned as these tend to be addressed primarily in terms of technical competencies underlying academic information skills (Weinert 2001), but according to Maybee and Zilinski (2015), data literacies can be better understood through a seven axis framework that includes: (1) Awareness: understanding data and its role in the society; (2) Access: understanding how to identify, locate and appropriately use structured data in datasets and databases; (3) Engagement: evaluate, analyse, organise and interpret data to make evidence-based decisions; (4) Management: organise and manage data; (5) Communication: synthesise and create visualisations and representations; (6) Data Ethics: identify diversified data sources, considering the risks of managing such data and the issues implicit in the use of data; (7) Preservation: awareness of long-term practices of storing, using and reusing data.

Critical data literacy approaches need to encompass how-to skills such as data analysis, curation, management, mining and visualisation (Gascó-Hernández et al. 2018) while also engaging with related critical, statistical, political, and media literacies, such as statistical, political, media and data. Statistical literacy is, for Wallman (1993), Schield (2004) and Mandinach and Gummer (2013), the ability to understand and use data effectively for decision-making.

For statistical literacy to be fruitful for citizenship, it needs to be complemented with political literacy, so people learn the basic principles and concepts of democracy and understand the functions and mechanisms of political institutions, as well as common legal language and concepts (De-Shalit 2004; Dudley and Gitelson 2010; Timmermann 2017; Wszalek 2017). To complement statistical and political literacies, citizens require media literacy to access, analyse, evaluate and create information in a wide variety of formats (Aufderheide 1993; Buckingham 2003; Jenkins 2006; Kahne, Lee, and Feezell 2012), and to critically evaluate the information presented in the press and social media networks (Culver and Jacobson 2012; Kellner and Share 2009; Littlejohn, Beetham, and McGill 2012). Furthermore, data literacies programmes need to include elements of critical thinking and research skills to foster public participation (Shirk et al. 2012), using open data in collaborative and multidisciplinary activities (Purwanto, Zuiderwijk, and Janssen 2018) to foster active participation and understanding of democratic issues at national and global level (Olssen 2004).

Davies (2010) states that in the future there will be a greater need for data capabilities in society to debate the meaning of data, and to find responsible ways of using open data. For Gurstein (2011), that the current trend towards public data transparency can create democratic opportunities to use open data for teaching and learning, which can help students to understand social problems and, according to Baack (2015), embrace the idea of open participation (Baack 2015). However, as Johnson (2014) notes, open data alone cannot promote social justice, as it can marginalise those who cannot engage with data effectively, rendering them objects of study.

A critical approach to data literacies must aim to empower educators and learners to be co-producers of knowledge (Andrade et al. 2011; Conole and Ehlers 2010; Cronin 2017; Orr, Rimini, and van Damme 2015), that can be supported with the use of open data to study authentic problems in society (Kasl, Marsiek, and Dechant 1997; Piorun et al. 2012), in a collaborative manner with a focus on evaluating information presented in a wide range of media and formats (Barron et al. 1998; Hmelo-Silver 2004) and to foster politically responsible decisions through the understanding of democratic and social processes (Ateas and Havemann 2019; Davies 2010). Strategies
to foster pedagogic innovation leveraging open data can enable them to participate as citizens (Baack 2015; Buttiglione and Reggi 2015; Huijboom and van den Broek 2011; Mellouli, Luna-Reyes, and Zhang 2014; Reggi and Dawes 2016).

Discussion and conclusions
The socially responsible role of HE in democracies must not only be to develop competencies for the labour market but also to be agents of social change, and most importantly, become open arenas to foster social participation through inculcation of transversal skills. Universities are key drivers to foster data literacies and must enable students and educators to challenge biased metrics, unethical uses of data, violations of privacy and the interaction of the datafied society with the quantified self.

Academic development programmes in data literacies need to include a wide range of perspectives, issues and challenges, as we have elucidated from the literature, including emergent components such as:

**Data ethics** can be understood as the morally responsible uses of data, caring for people and society adhering to the principles of human rights and personal data protection.

**Data politics** encompasses the political aspects of data including ways in which data are collected, accessed and displayed by the governments, and how these data are used to foster participation and policy making.

**Data governance** can be understood as the policies and regulations in place for deploying and presenting data in regards with its accessibility, usability, integrity and security-based data standards, norms and laws.

**Data management** is the process of acquiring, validating, storing, protecting and processing data, in order to ensure its accessibility, usability, integrity, security, reliability and timeliness, and to enable users to search, retrieve, appraise, assess, mine, prepare and clean for analysis.

**Data analysis** is the processes and methods in which data are collected, organised, assessed and studied to obtain and extract useful information from it.

**Data narratives** are methods such as data journalism techniques and data storytelling that provide an innovative yet rigorous way to communicate research beyond scientific and technical endeavours in an accessible way to inform wider audiences.

**Data visualisation** is the process of transforming data into visual models to make information understandable and informing to support decision-making or comprehension of processes and phenomena.

These elements have the potential to (1) seek to embed political, media and statistical literacies and to develop transversal skills for lifelong learning; (2) enable students to understand and critically analyse information and data from media and government sources; (3) situate learners as critically engaged data intermediaries who are empowered to act as social detectors and effectors in the service of social justice and democratic values.

Owing to their dual education and research mission, universities and academics are well placed among social institutions to support young citizens to develop the necessary critical literacies to ensure widespread democratic participation and social
justice (Timmermann 2018); but while our focus here is on the role of HE in fostering literacies for democratic citizenship, it is important to keep in mind that universities are only one of the many groups of social actors that need to be mobilised for such efforts. Educators should also consider how they and their students might engage wider audiences in such debates through OEP.

A true democracy demands universal participation, and active efforts need to be made to tackle the problems of discrimination and adaptive preferences; therefore, bridging the data divide requires substantial efforts to reach all citizens (Schoonmaker 2018; Warschauer 2011). Working towards social justice will be a lengthy endeavour, but one place universities and educators can start is with critical pedagogies and curricular design, which give students the tools to disentangle the processes, values and power behind data.

References


Gurstein, M. B. (2011) ‘Open data: empowering the empowered or effective data use for everyone?’, *First Monday*, vol. 16. doi: 10.5210/fm.v16i2.3316


Citation: Research in Learning Technology 2020, 28, 2468 - http://dx.doi.org/10.25304/rlt.28.2468


