

PRE-PRINT – Please cite as:

Raffaghelli, J.E. (2022) Educators' data literacy: understanding the bigger picture. In Pangrazio, L. and Sefton-Green, J. "Learning to Live with Datafication. Educational Case Studies and Initiatives from Across the World". Routledge. DOI: 10.4324/9781003136842-5

# Educators' data literacy: understanding the bigger picture

Juliana E. Raffaghelli – [jraffaghelli@uoc.edu](mailto:jraffaghelli@uoc.edu)

## 1- Introduction

Data extraction and algorithmic manipulation have become increasingly frequent across the globe – hence the expression “datafied society”. Initially perceived as an opportunity to innovate in several areas of human knowledge (Stephen-Davidowitz, 2017), data practices had unthinkable impacts. Race, gender and other vulnerable characteristics were made invisible, overrepresented and over-tracked by certain forms of data visualisation (Ricourte, 2019). Automations based on machine learning have entailed perilous biases in the way they represent or neglect relevant cultural perspectives (Malik, 2020). Moreover, they have led to biases and users' lack of agency or even harm (Eubanks, 2018; O'Neil, 2016). In education, research has also highlighted forms of surveillance on children, teenagers and young adults with the aim of addressing their behaviours, emotions and cognition (Chi, Jeng, Acker, & Bowler, 2018; Lupton & Williamson, 2017; Prinsloo, 2020). The “platformisation” of learning and the monetisation of students' data are also pressing issues in the educational agenda. These problems were magnified by the intense and unprecedented use of digital technologies during the pandemic (Perrotta, Gulson, Williamson, & Witzemberger, 2020; Williamson, Eynon, & Potter, 2020).

Against this backdrop, an activist counterculture emerged (Baack, 2015) in which participants aim to uncover how surveillance is redistributing power through the participatory appropriation of data (Lehtiniemi & Ruckenstein, 2019). Expressions of this movement have manifest through the Open Government Data and Open Science (Lehtiniemi & Ruckenstein, 2019), as well as with independent collectives reacting to the oppression of surveillance, including forms of disconnection and “hacking” the system (Pybus, Coté, & Blanke, 2015). Overall, data activists' search for recognition, awareness, redistribution of symbolic and material power aims at emancipation.

As Milan and van der Velden (2016) argue, data circulates generating imaginaries, discourses and practices, which they refer to as “data epistemologies”. Despite the universal circulation of metrics and quantification, the concept of data epistemology spots the contextual and diversified nature of data practices in response or reaction to datafication. According to these same authors, the activists that search for data appropriation to express diversity and empowerment embrace a proactive data epistemology . Instead, the activists whose goal is to uncover injustice and condemn malpractice stick to a reactive data epistemology. In each case, the quest for equal representation, voice and self-determination leads to an ideal known as fair data culture The term “epistemology” refers too widely to a theory of knowledge in the context of philosophy, though it is apparently used “lato sensu” by Milan and van der Velden. From an

educational perspective, my point here is the need to find an “actionable” concept that allows us to move from beliefs, attitudes and somehow unconscious representations of data which trigger human activity in general, and the educators’ practices in particular. In this regard, we could adopt “dispositions toward data”, a concept that signals the attitudes and beliefs towards data and therefore how people might react to it. Indeed, the dispositions can be seen as a continuum, a vector which extremes are represented by proactive and reactive ideas, practices, discourses and responses to data. Nonetheless, transforming both reactive and proactive dispositions toward data into action require advanced technical skills and civic engagement that go well beyond the actual possibilities marginalised and vulnerable collectives may have.

In this endeavour, education plays a crucial role by developing data literacies that move between proactive and reactive dispositions, thereby finding a balance between these two poles. Moreover, educators’ awareness of their own dispositions toward data might have impact in the way they move to pedagogical practice embracing the complexity of data. This statement, however, has immediate implications: if education has to respond to the context of innovation and continuous transformation of data practices, and if several data approaches can be embraced, the educators’ response could become fragmented. As a result, it would be difficult to promote creative and agentic ways of relating with data in education and the society (Raffaghelli, 2020b). In fact, the initial research around data literacy has focused a narrow and instrumental conception, which does not promote complexity in data practices (Raffaghelli & Stewart, 2020).

In this chapter, I contend that there is need of a more complex perspective about data in the society and in education, in order to address holistic, integrated and critical educators’ professional practice. Indeed, there might not be a universal perspective on educators’ data literacy that institutions and the system can embrace. Instead, my claim is that educators and institutions should engage in explorative approaches unveiling diversified dispositions to data. To support this assumption, I will introduce the results of 12 workshops (8 with English speaking participants, 3 with the Spanish speaking community and 1 in the Italian context) which I conducted from 2018 to 2020. Overall, nearly 294 amongst educators, PhD students and faculty professors participated. The methodology was held constant along the several workshops: I introduced several data practices in education and society, to trigger reflection about the educators’ professional knowledge and practice. Moreover, I considered the workshops as part of hermeneutic circles where the conversation between me and the participants triggered understanding around educators’ literacies to deal with situated data narratives and practices. This conversation unveiled the tensions between dispositions to data (reactive or proactive) and the forms of access to data (public or private), as spaces where the educators intervene. The data literacies required are diversified across these spaces, but they are also part of an integrated mosaic that can vary from a context to other. My ultimate consideration is that faculty, students and other stakeholders’ efforts to develop data policies, strategies and practices, should not overlook the “bigger picture”

## 2- Conceptual Lens

Promoting complex understandings of data in society requires embracing approaches that support learners to become aware of and actively take control over data practices, both as individuals and collectives. (Markham, 2018). While it is clear that technical literacies are relevant to read, manage, process and visualise data, and interact with algorithms and develop a proactive disposition to data, critical literacies around a participatory and dialogic culture explore meaning, power and the material conditions of data practices developing reactive disposition to data. In this sense, a complex understanding of data literacy

goes beyond the technical and the critical component and encompasses the integration of the data approaches. In the following, I will consider some theories and concepts helping us to understand the evolution of data literacy as research topic.

Paulo Freire's critical pedagogy has already tackled the need of integration between the technical and the socio-cultural layers (Freire, 2000). Through his theoretical lens, the transfer of knowledge from the *élite* to the community is to be overcome and redirected towards an approach where the community builds contextual knowledge by appropriation of the technical means in order to represent such world. For Freire, these means were reading and writing. Later on, the digital society shifted the debate on the digital medium and the needed skills. While there has been strong attention to the technical skills required to be digitally literate, the ability, the knowledge and the dispositions to live in a digital society, are also necessary. Using digital technologies critically also requires some degree of designing and creativity (Pangrazio, 2016). In short, technical and critical skills are important to engaging with the digital medium in a way that is both safe and creative (Tour, Creely, & Waterhouse, 2021). Moreover, as Pöttsch points out, digital literacies need to move beyond responsible and efficient digital technology use and focus on 'capitalist dynamics, environmental ramifications, and individual empowerment' (2019, p. 222).

Nonetheless, in recent years much emphasis has been placed on coding (Popat & Starkey, 2019), the maker movement (Papavlasopoulou, Giannakos, & Jaccheri, 2017) and digital competence (Carretero, Vuorikari, & Punie, 2017). The initial vaunted-potential of digital competence to be part of a rapidly changing and smart society turned later into a series of predictable failures to promote access or socially relevant learning, beyond the technological skills (Sancho-Gil, Rivera-Vargas, & Miño-Puigcercós, 2020).

Similarly, the debate around data literacy focused initially on the technical dimension (Raffaghelli & Stewart, 2020). The discussions on numeracy and then on statistical literacy paved the way for a technical perspective on data literacy (Gould, 2017). This strand of research addressed the technical comprehension of algorithms, data-driven practices, data visualisation and programming as the creative side of data science (Gray, Gerlitz, & Bounegru, 2018; Wilkerson & Laina, 2017). Aligning with this trend, the digital tools to process and visualise data, as well as the opportunities of triggering dynamic representations, recommendations or activities, opened an area of skill development in higher education. Overall, data literacy mainly covered individual practices within the classroom, both from the side of data as educational content and data as part of inquiry tasks (Maybee & Zilinski, 2015). Nevertheless, the aggregation and analysis of data gathered from assessment and processed as quality evaluation metrics can initiate reflection, with data considered as an element in support of pedagogical methods (Mandinach & Gummer, 2016). Less emphasis has been put on understanding how the students can be empowered by the use of data within the classroom and towards their communities (Raffaghelli, 2020a), though the debate around students' data usage is evolving at a steady pace (Willis, Slade, & Prinsloo, 2016). More recently, ideas such as data justice and data feminism have permeated the educational system in search of broader perspectives on data for the society and the educational institutions themselves (D'Ignazio & Bhargava, 2015; Raffaghelli, 2020b). Nonetheless, authors have advanced the need of approaching data literacy through a critical lens (Pangrazio & Selwyn, 2019), as well as through the adoption of Freire's perspective (Manca, Atenas, Ciociola, & Nascimbeni, 2017; Tygel & Kirsch, 2016).

Moving beyond the technical approach, critical data literacies can also explore different data practices: from promoting open data appropriation in formal and informal contexts of learning (Baack, 2015; Manca et al., 2017) to considering personal data privacy (Pangrazio & Selwyn, 2020), awareness of data usage and algorithmic manipulation across digital learning platforms (Stewart, 2020).

The literature considered here highlights the problem of fragmentation, and the tensions between different dispositions toward data re-emerges and produces entropy in the way the educational intervention might be characterised and implemented. The liaisons between the several data usage scenarios and which

approach to take is left to the educator who is often in professional isolation. To strike a balance between the proactive and reactive dispositions toward data mentioned in the introduction, the emerging data practices should be identified and conceptualised to support educators' reflection and organic action, as well as to promote critical data literacies through a complex integrated data literacy's "bigger picture".

How could be the educator's isolation overcome? In the following, I will introduce my approach, based on building professional learning spaces where to reflect about the complexity of data in the society and within the educational practice.

- 3- Methodological approach: a hermeneutic journey to explore the boundaries of data literacy with and for educators.

### *3.1. Context of implementation and research goals*

In the context of a five-year (2018-23) national project supporting Faculty Development in changing digital contexts<sup>1</sup>, my overall research work aimed at exploring and mapping data practices in Higher Education Institutions (first and second year) as means to build a framework (third year) that would be helpful for educators to promote critical data literacies. During the initial phases of the project, the map of data practices proved to be fragmented, mostly connected to a discourse of innovation (e.g., learning analytics) and to open science (e.g., open data) (Raffaghelli, 2020b). For the scholars and the institutions I interacted with data was seen as something happening in the present but continuously evolving to shape either a utopic or dystopic future: If we do not catch up data-driven practices, will we step outside a trend of necessary innovations to promote new skills and new jobs? Will data practices produce more harm than good? Based on an interdisciplinary approach, the project promoted a vision of data literacy that included technologies, politics, aesthetics and narratives of data, searching for social justice as the ultimate goal of any educational action (Raffaghelli, 2018).

In this context, I carried out a number of workshops for educators, as part of an initial exploratory activity to understand what actions were needed to build educational practices focusing on data in the society. The educators and educational researchers' specific goal was to identify the impact of emerging data discourses and practices in connection with their professional identity. Most of the workshops were conducted during international conferences, though some were part of faculty development strategies.

### *3.2. Research and educational design*

From a methodological point of view, the workshops could be characterised as nodes in an iterative action conducted by the researcher with educators' engagement, aimed at building a bigger, integrated picture around data practices and the needed literacies. For this purpose, I worked on the philosophical concept of the "hermeneutic circle" (Gadamer, 1960/2013). The German philosopher conceived it as "a movement" in "repeated return from the whole to the parts and vice versa" (p.196), between the reader and text, within conversation. Hence, it involves a dialogue with the world and others (Gadamer, 1960/2013, p.386) and refers to a situated knowledge, prompting a contextualised search of truth, which cannot be reduced to a set of research methods. This approach has already been embraced in educational research – i.e. in-depth

---

<sup>1</sup> Professional learning ecologies for digital scholarship: modernising higher education by supporting professionalism – Ramón y Cajal Programme RYC-2016-19589 (<http://edulab.uoc.edu/en/projects/led-projects/professional-learning-ecologies-for-digital-scholarship-modernizing-higher-education-by-supporting-professionalism/>).

interpretation, jointly with the learners and their communities – leading educators to frame educational problems, design for learning and search for adjustments between the educational act and its impact on the reality (Carr, 2007). In this study, I developed a workshop design with activities and a conceptual tool (listed below) to think on data practices, which became the participants’ object of critique and enrichment. This engagement elicited feedback and challenged our (educators’ and mine) understanding of data literacy as a whole, exploring the spaces for educational intervention. In this regard, the sampling was theoretical, not based on the participants and their responses, but on the categories of analysis’ saturation (Corbin & Strauss, 2007).

Since the topics tackled different data practices, the workshop structure as a hermeneutic-pedagogical approach was constant and constituted of:

- **Initial information** to generate a sense of curiosity;
- **Self-diagnosis** to capture insights and develop self-awareness;
- **Technical instruments to engage with data as material expression**, based on methods, jargon, technologies, practices and knowledge within professional communities;
- **Reflection** connected with the participants’ experiences and practice contexts based on the epistemologies of data. *‘What am I doing and why am I not working with data? Why am I engaged in this movement, or why am I not? Am I resisting the establishment, or am I trying survive within an unjust context of professional practice and personal life?’*
- **Exploration of the professional roles as educators and educational researchers:** *‘if I behave in a certain way, to what extent am I providing a different vision to my students or to the participants of my educational research?’*
- **Representation of the reflection’s outcomes through a quadrant graph as a conceptual tool.** Elaborating on each of the conversations with the educators, I modified the pair of tensions forming the quadrant. The dimensions generated spaces to help categorising the educators’ fragmented, lived, perhaps untold, experience. Figure 1 introduces the initial quadrant graph of “data practices and literacies”, being the goal to improve the categories and to characterise the types of data literacies needed to move across the four quadrants. In each of the four quadrants, spaces are delimited by the tension between access to data/knowledge and the educators’/learners’ dispositions toward data. The theoretical foundations were discussed previously, considering public data access and the open data movement, as well as private data access and the concerns of data surveillance and algorithmic manipulation. During the workshops, my aim was to place a series of specific data practices within each area, followed by the required educational practices to promote data literacy. A recursively complex picture around data practices and literacies could have been achieved through the educators’ participation, reflection and understanding across the four fields.

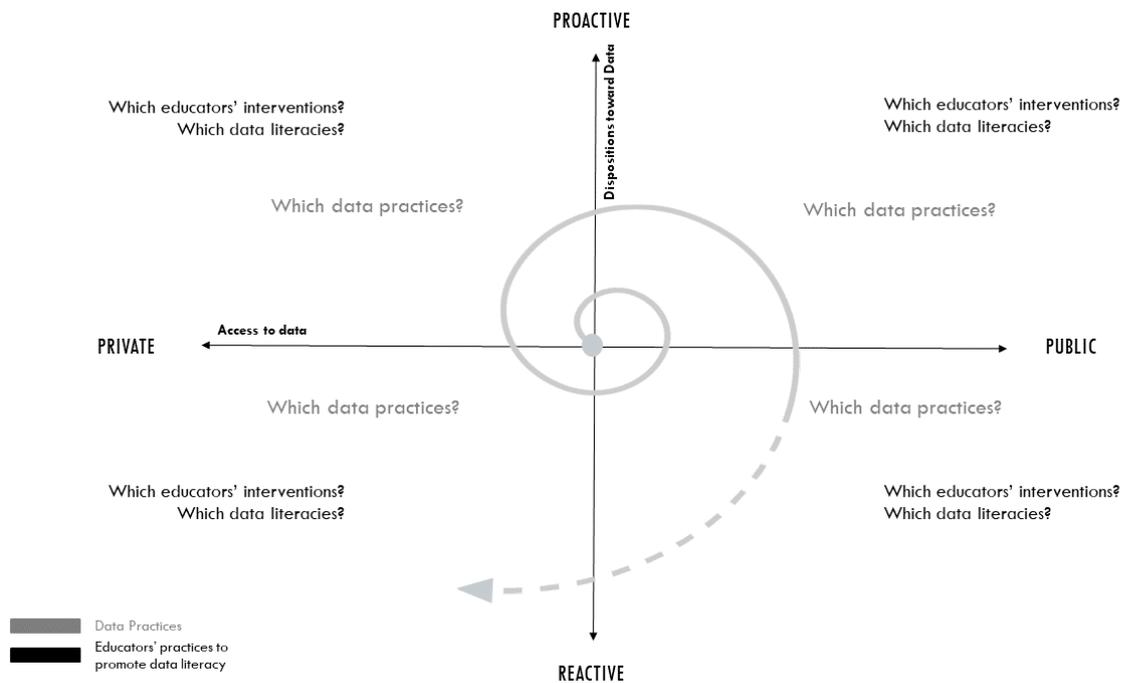


Fig. 1 - The quadrant graph as conceptual tool to accommodate educators' perspectives on data literacy around data practices.

Table 1 shows the workshops' names and link to the published open resources and reports. Each workshop could be considered the beginning and the end of a hermeneutic circle. The table also includes the number of participants, the working languages, the participants' expertise, the expected targets, the number of encounters and duration. Also, the geographical areas covered can be found in the seventh column. All the mentioned columns portray the extension and diversity of education professionals who participated and engaged in the activity. A breakdown point generated by the beginning of the pandemic was noted, so I moved the organisation of the workshops to online activities.

The most relevant element, that is, the workshop focus on data practices, is reported in the "data approach" column. The data approach shaped the topics dealt with and the types of resources and activities (though the methodology was constant). In that regard, I selected the workshop topics over the prior hermeneutic iteration.

The data approaches adopted in each workshop were:

- A= Open Research Data for Open Educational Science;
- B= Open Data as Open Educational Resources;
- C= Data Surveillance in Education and Society;
- D= Data for good in education.

Workshop's Name	Sched.	N	Lang	Expertise	Targets	Geographical Distribution	Time	M	Data Approach	Mode	Report
W1/ NWL2018	May 2018	9	ENG	Advanced	Educational Researchers Higher Education	Europe	60	1	A	onsite	<a href="http://doi.org/10.5281/zenodo.4446013">http://doi.org/10.5281/zenodo.4446013</a>
W2/ EDEN18	June 2018	8	ENG	Advanced	Transnational Projects' Leaders Educational Researchers VET educators Open Education promoters	Europe	90	1	B	onsite	<a href="http://doi.org/10.5281/zenodo.4446267">http://doi.org/10.5281/zenodo.4446267</a>
W3/ MASTERUB19	Feb 2019	16	SPA	Mid-career Advanced	Teachers' Educators, Educational Technologists	Europe Nord America Latin America	240	2	B	onsite	<a href="http://doi.org/10.5281/zenodo.4446763">http://doi.org/10.5281/zenodo.4446763</a>
W4/ JTELSS19	June 2019	25	ENG	Initial	PhD candidates in technology-enhanced learning	Europe	60	1	A	onsite	<a href="http://doi.org/10.5281/zenodo.4446844">http://doi.org/10.5281/zenodo.4446844</a>
W5/ OEG19	Nov 2019	6	ENG	Mid-career Advanced	Transnational Projects' Leaders Educational Researchers Open Education promoters Higher Education experts	Europe Nord America	60	1	A-B	onsite	<a href="http://doi.org/10.5281/zenodo.3557431">http://doi.org/10.5281/zenodo.3557431</a>
W6/ UOC2019	Dec 2019	12	ENG	Initial	PhD candidates Research Assistants in technology-enhanced learning	Europe Nord America	180	1	C	onsite	<a href="http://doi.org/10.5281/zenodo.4453078">http://doi.org/10.5281/zenodo.4453078</a>
W7/ UBAAP2020	Jan 2020	10	SPA	Mid-career Advanced	Higher Education Teachers	Catalonia-Spain (Europe)	180	1	D	blended	<a href="http://doi.org/10.5281/zenodo.4460680">http://doi.org/10.5281/zenodo.4460680</a>
W8/ OER2020.I	April 2020	73	ENG	Mid-career Advanced	Transnational Projects' Leaders Educational Researchers Open Education promoters Higher Education experts	Europe Nord America Latin America	30	1	B	online	<a href="http://doi.org/10.5281/zenodo.3739180">http://doi.org/10.5281/zenodo.3739180</a>
W9/ OER2020.II	April 2020	94	ENG	Mid-career Advanced	Transnational Projects' Leaders Educational Researchers Open Education promoters Higher Education experts	Europe Nord America	60	1	C	online	<a href="http://doi.org/10.5281/zenodo.3744135">http://doi.org/10.5281/zenodo.3744135</a>
W10/ AREA2020	April 2020	25	SPA	Mid-career Advanced	Educational technologists Open Education promoters Higher Education teachers Teachers' Educators	Latin America	90	1	B	online	<a href="http://doi.org/10.5281/zenodo.3839439">http://doi.org/10.5281/zenodo.3839439</a>
W11/ MED2020	Nov 2020	5	ITA	Initial Mid-Career	Teachers' Educators, Educational Technologists	Italy (Europe)	240	4	B & C	online	<a href="https://www.youtube.com/watch?v=F0Z-BWGPcwQ">https://www.youtube.com/watch?v=F0Z-BWGPcwQ</a>
W12/ LATAM2020	Dec 2020	15	SPA	All levels	Higher Education Teachers	Latin America	60	2	A to D	online	<a href="http://doi.org/10.5281/zenodo.4475965">http://doi.org/10.5281/zenodo.4475965</a>
<b>TOTAL WORKSHOPS: 12 - TOTAL PARTICIPANTS= 298</b>											

Table 1 – Implemented Workshops

#### 4- Results: Interpreting and understanding data practices with the educators

As observed in table 1, data practices were concentrated mainly in the initial workshops (W1 to W4), which began with open data in educational research and moved to the use of open data as open educational resources, sticking to epistemologies of proactive data. The educators' participation and questions were oriented towards the exploration of ethical concerns and possible implications for the educational practice from the perspective of epistemologies of reactive data. The starting point was based on data practices in education that I had explored in my prior research. However, I purposely provided an incomplete structure with an aim to trigger the hermeneutic cycle and complete the quadrant graph as a final representation of the collective understanding of educators' engagement with data literacy in their professional activities. In the succession of workshops, I either invited experts or collaborated with colleagues whose expertise would support participants' reflections while I led the workshops' methodological approach. In the following, I introduce quotations representing key elements in some of the most relevant hermeneutic iterations.<sup>2</sup>

Initially, I used the quadrants from the first circle in the A-B type workshops to characterise the personal motivation versus the institutional support to become open scholars and share/re-use open data. By exploring the several positions and beliefs around the role of data in society and the educator's profession, I became aware that the dispositions toward data played a crucial role. As it emerged in the W1 (Report W1<sup>3</sup>, p. 3), some participants (SJ and PM) were concerned on how data could be extracted whereas MH considered the ethical conundrum in using students' data:

*"If a person or student leaves their data on a social network, even if it is public, can it be collected and processed for research? This was not the original purpose (MH)"... "one bigger problem is the data that could not be accessed due to the obscure forms of data processing by social networks" (SJ). "It appears that data, which can be necessary for research purposes (and hence for good, hopefully!), is more and more difficult to get" (PM). "...data is more and more difficult to get (due to ethical concerns and procedures) whereas for the companies data is easily available" (SJ).*

In the following workshops (W2; W3), the problem of open data extraction or usage within the classroom for teaching purposes was considered by the participants. I continued exploring the right side of the quadrants. Educators held that providing the class with open data as resources, as well as promoting civic engagement and monitoring open data could imply innovation and discovery, hence, massive intellectual efforts (Report W2, p.8):

*"The main issue was to identify which data and which levels of literacy had to be considered: my students would not be able of participating in open hackathons" (FN). "The materials should be produced to prompt relevant reflections and engagement with data, whilst also being aware of the risk of failure due to the technical complexity of extracting, managing, elaborating and presenting data (to students)" (SM)*

On these grounds, for the hermeneutic circle referring to OEG19 (Report cited in the Table, p.4), I re-proposed the quadrants to explore public vs private knowledge and personal vs group generated data in order to better grasp the expressed problems of access to data. The participants (divided in two groups) engaged with these categories in the following ways:

*"Awareness and debate come first (relating data practices)! But we are not sure about the policy making that could come at a late stage" (W5-NS.Group1; "We need training in educational data management &*

---

<sup>2</sup> The participants' profiles (after anonymisation procedures) quoted in this paragraph are presented in the annex I

<sup>3</sup> In this section, the reports cited are listed in the Table 1

*and research data management...know the basics before you dive in deeper"* (P3.G1). However P5 insisted on the relevance of AI and data *"Awareness and Debate to support participatory design for Artificial Intelligence Integration will be at the cutting edge"* (W5LH.Group2). Within this group, the participants concluded indeed that the category of "data analysis and algorithms transparency" (Report W5, p.4) was considered transversal to other forms of knowledge and awareness.

Again, the different approaches between groups in this workshop highlighted the presence of diversified dispositions to data. The attention in this group was all devoted to the need of making data a public good (Report W5, p.5).

From the sixth workshop, the focus moved completely to data practices connected to surveillance. The quadrant graph illustrated the pair reactive/proactive, but was mostly centred on (private) data extraction. In this case, the educators moved again from one side to another expressing their diversified positions about data usage (W6, p.11)

*Researcher: "I think datafication in higher education is... (open ended statement completed through written post-its)*

*DM: A Bell curve: what really matters (dark for data and eye)*

*PD: Data? Useful*

*AK: A way to decrease education funds and increase the gap between privileged and non-privileged*

*WW: Data for the future: To determine the data present in different activities and use it with educational purposes*

*MP: ...is a two-sided blade*

*CB: ...a way to kill the art of education*

Most importantly, in this workshop the relevance of technological skills was highlighted as means of preparing for proactive "hacking" of AI biases. While talking about Alexa, one female "tech-savvy" participant said: *"...well actually I have another example because in my house we programmed a minicomputer -with my room-mates- to make the voice assistant, but we chose a male voice, and we called it 'Ramón' (Report W6, p.13). Indeed, this group discussed the need of a proactive approach in education to develop the skills as "data hackers" in search for public and private data appropriation.*

From the W7 cycle, I kept the perspectives based on public vs private or open vs restricted access and provided the pair of tensions constituted of the reactive (defensive towards data that harms) and the proactive (fostering data practices, even in the case of hacking). This workshop was mainly focused on data in the educational process (particularly learning analytics). The questions raised led us to understand that data practices move from educational research to teaching practice, from the open knowledge as a scholarly endeavour to the protection of students' data, and that more, better and bigger data does not equate to quality research, teaching and learning. Nevertheless, the participants (all coming from the Spanish context) were mostly reliant on the use of data to inform their practices at the beginning of the process (Report W7, p.10). By the end of the process and during the pandemic, participants were mostly concerned with the failures of the system (*unavailable data is a problem..., like for example when you use blended approaches, W7.LP*). However, against the pandemic situation where all the activities were transferred to the online learning environment, they also concluded that *"more data does not equate to more knowledge or an optimal reading of reality"*, but *"good data, shared amongst teachers, could lead to a better educational practice, W7.EP"*. This circle contributed to the understanding of educators' professional knowledge on educational data, from the private sphere (students' data care) to the public (as a source of educational reflection).

In the following cycles (from W8 to W12), the quadrant graph remained stable, but the participants' engagement helped to saturate or better inform the relevant categories. For example, at the W9, the participants expressed the need to raise (students and teachers') awareness on data practices based on the familiarisation with technology in educational settings (Report W9, p.12):

*"(we as educators should) let them (students) know that everyone is on the same learning curve and that we are all at the same time experiencing something new (data surveillance). We can improve teaching digital literacies to students if we also improve our understanding of digital literacies (AT)"*

During both W8 and W10, the need to engage students as data creators and collaborators was also considered. The following conversation expresses this need (Report W8, p.14):

*ML: Would it be interesting to create student-researcher data communities?*

*Researcher: You can all create your community for your class and put the data and resources there*

*AMS: I would love to help!*

Above all, at the W11 and the W12, I was able to introduce a more complex picture of data practices. I worked together with the participants of W11 in the elaboration of learning units. At the W11, participants worked integrating technical data skills required to navigate and extract data. They used skills to interpret and criticise data visualisations, and they also created data sculptures and data storytelling. At the W12, I provided technical instruments to understand how Higher Education collect and use students' data, highlighting the limitations of many approaches. I also emphasised the possibility of opening spaces for resistance to detach from dominant discourses about data and build more tailored approaches within the digital classroom. The participants' response again supported the fact that they held different dispositions toward data (seemingly closer to those found in the W7 case). Many participants attended the workshop to collect information on "innovations", others were interested in expanding their technical knowledge. However, there was also a group concerned with data sovereignty, connected to local and regional debates around Open Source software in education. This group of participants was interested in curating their data as teachers and researchers, avoiding "academic colonialism" when adopting well-known platforms to teach and to do research:

*HPG: in my house I created a Moodle server and Big-Blue-Button where I have all my students' data and it is private (Report W12, p.12).*

*YMF: What to do when data access is not collectivised and those who have it are not the most suitable to guide an adequate policy? (idem, p.21)*

*CPB: What do you think of academic capitalism? Is research (no matter how academic) mediated by economic interests and even personal groups? How to get rid of it? (idem, p.21)*

As a result, the conceptual scheme based on the quadrants was enriched as shown in the Fig. 2

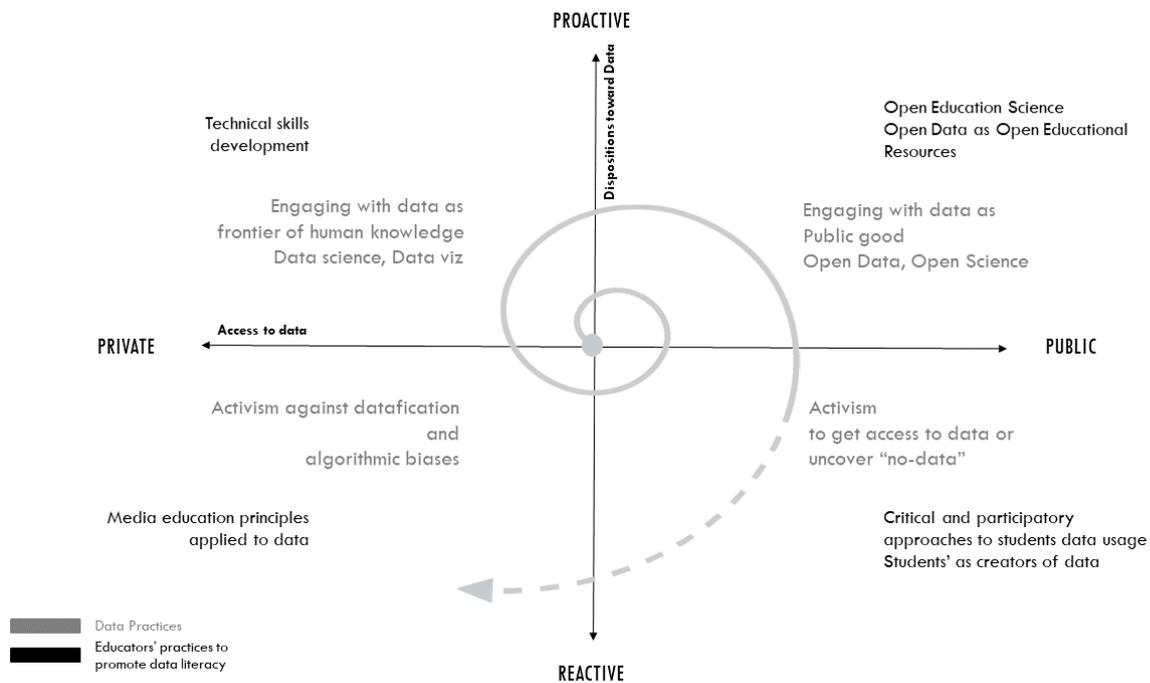


Fig.2 – The results of the iterative process along the workshops through the scheme of the quadrant graph as a complex integrated picture on data practices & literacies.

### 5- Discussion: educators' data literacy through a complex lens

In this section, as captured in the quadrant's graph, I will discuss the evolution and conceptualisation of data practices in education and society, as well as the possible data literacies educators could engage in. This is the main result after the iterative hermeneutic circles along with twelve workshops devoted to educators professional learning.

The first quadrant covered the phenomena of open data as a public good and it was based mainly on the results of the W1 and W4. The reflection arising from it supported the relevance of proactive dispositions toward data as part of the open educational research practice, though some participants were wary about students' data usage. In this case, the required literacies would be browsing open research data, as well as using and generating it. It is worth considering that the literature reports the need to support the open data movement in research as part of Open Science and open access to public knowledge (Borgman et al., 2015). As expressed in the introduction, the proactive dispositions toward data support the idea that data should be used by others, namely, students, scientific communities and citizens (van der Zee & Reich, 2018; Wilkinson et al., 2016). Nonetheless, this is still an issue for most research areas, particularly for education (Raffaghelli & Manca, 2019). Therefore, educational researchers and educators should cultivate this literacy to support fair (open) data practices. However, the participants were concerned about the difficulties of accessing the private platforms to unleash data to researchers (particularly social media, but also learning management systems). This was deemed an obstacle to furthering open science in education, though it opened to the exploration of a different set of literacies connected with platforms and the unfair use of data.

Moving to the second quadrant (activism to get access to data or uncover no data), the hermeneutic circles (W2, W3, W9, W10) focused on open data usage into the pedagogical practice. The participants' engagement with the topics of the upper and lower quadrants enriched the idea that open data usage in

society, open data in education for a teaching/learning scholarship, and open data as open educational resources require careful collaboration, institutional support, as well as initial training that help educators to understand data quality and the way data might be presented to the students. Nonetheless, in these groups the focus was more on technology usage than its critique.

These two circles highlighted the need to promote awareness-raising on reactive dispositions toward data, concerned about the abuse in the usage of private data. In the pandemic aftermath, there was an increasing debate around the platformisation and monetisation of education (Williamson et al., 2020). However, by 2019, when the initial bunch of workshops were implemented, this was not “*vox populi*”. In any case, the OEG2020.I and the AREA2020, along with the geopolitical representation of Latin American participants, yielded some voices around participatory approaches and a more democratic usage of data. Indeed, the perspective evolved about how the students might play a role in supporting marginalised populations to collect, interpret and tell their own data stories, even for social problems for which data does not exist. As argued by D’Ignazio & Klein (2020), intersectional data practices face data misrepresentation and overrepresentation of some collectives, in search for co-liberation, namely, the possibility to appropriate of advanced data science to support agentic practices. This is a fact that educators should consider when designing for learning with open data, moving from the upper to the lower part of the quadrant’s right side.

Moving to the third quadrant (activism against surveillance and algorithmic biases), the W6 circle clearly detected data surveillance. Indeed, the participants’ understandings and practices on data surveillance (both active or passive roles) were evident. I noticed that those with advanced technological skills felt their practices as more “agentic” and empowered to resist, hack or even generate alternative technology development and usage. Aligning with critical pedagogy, technological understanding (or basic literacy) should not be an instrument of adaptation to the system but of resistance and activism (Tygel & Kirsch, 2016). Across the majority of workshops, those with less technical skills pointed out that educators are not “algorithms makers”, and they are left with little space for intervening in what they defined as an overwhelming acceptance of datafication. This finding was coherent with Pangrazio & Selwyn (2019). However, as shown in the W7, educators felt confident enough to protect their students’ interests, although this action could only cover the educational spaces controlled by the educator, and probably could not be extended to the several spaces connected to students’ lives.

In this sense, the educators’ professional development also requires data literacy around the surveillance systems, data ethics and alternative software protecting students’ data. As it has been pointed out, this relates to the software they work with as educators (Wasson, Hansen, & Netteland, 2016), but also to surveillance, which in society has become a topic that can be developed in class as part of familiarisation with technology (Stewart, 2020) and media education (Hobbs, 2020). During the pandemic (November 2020), the W11 circle supported this reflection as well. The participants’ attention towards the problem of datafication and data practices was very high at that point since it was also connected with the so called “COVID19 infodemic” and the consequences of ill social media usages.

This circle highlighted the need to enter into the fourth quadrant (the engagement with data as a frontier of knowledge), and it was also supported by the findings in the W12 workshop: for educators, in order to embrace a complex approach to data practices, and to intervene accordingly in the development of the data literacies needed, the technical awareness and creativity must be put in tension with the political and ethical concerns. Building on Bhargava and D’Ignazio’s concept of popular data literacy (D’Ignazio & Bhargava, 2015), the distribution of symbolic power and material resources through education requires the educators’ interaction with institutions to build situated approaches that help understanding data and how to use it creatively.

Nevertheless, the quadrant graph would also allow the educator to enter from one or the other side, from the lower or the upper part, covering what is needed at a particular point of the curriculum or from the specific subject field. The quadrant, hence, can be an approach to explore and integrate different forms of data literacy.

## 6- Conclusions

In this chapter, I argued that data as social, political, technical and educational problem is complex and evolving. I contend that a situated and explorative approach to understanding data should be considered. This culminated in a conceptual tool built with 292 educators engaged in twelve workshops, to unveil data practices' complexity and facilitate further educators' professional interventions in the promotion of data literacy. This process of understanding the data landscape in search for the bigger picture yielded four types of social activity around data, as experienced by the educators and me: open data as a public good, activism to access existing data or uncover misrepresentation (no data), activism against dark datafication and algorithmic biases, and the engagement with data as a frontier of knowledge. The literacies needed in each of the quadrant's spaces were specific, developed over time and related to the faculty, students and other stakeholders' efforts to make sense of the several dispositions toward data as a base for data politics and practices (Raffaghelli, Manca, Stewart, Prinsloo, & Sangrà, 2020). However, such literacies cannot be seen as separated pieces, but rather, as a complex puzzle.

As the reader might presume, the institutional scenarios can be various. Discourses on the promising logic of Big and Open Data could be heralded by one institution, whereas in others there could be complete silence and lack of policies around the issue. In the best-case scenario, there might be a participatory approach towards the exploration of the problem to generate guidelines, policies, manifestos or strategic interventions in order to deal with datafication and surveillance, as well as with data for empowerment and co-creation (Raffaghelli, 2020b). Overall, data practices do not relate only to the availability of technological devices or the connectivity across networks, and neither to data processing performance in terms of volume and time. Data practices instead might be wired to the several groups' dispositions to data.

By exploring the four quadrants proposed here, a group or an institution might be prepared to plan a more strategic professional development for educators. The quadrants could address planning, decision-making, interventions and even activism relating to institutional and individual choices on the digital ecosystem and its consequences. This might also disentangle personal, professional and collective dispositions toward data.

Is data good or bad for our societies, and for our educational activity? Just as our interpretive journey through data practices does, this is a question that requires a holistic, complex perspective. This means to build integrated pathways towards data literacy as situated learning about data practices. To this end, educators and educational researchers should collaborate with policymakers, activists, the private sector and society, building commitment towards data-driven designs, processes and data products connected to the topics taught and within the educational institution. Concretely, it means to have a deep understanding of what the workflow of data-driven, algorithm-design, coding, data visualisation (proactive data epistemology, both private and public access to data) implies, as well as of the institutional and social impacts of data practices (reactive data epistemologies, both private and public access to data).

In Gadamer's terms, "we learn a *techne* and can also forget it. But we do not learn a moral knowledge, nor can we forget it" (1960/2013, p.327). Hermeneutical consciousness (conceived as understanding leading the right action) "is involved neither with technical nor moral knowledge, but these two types of knowledge" (p.325). Indeed, data practices require a technical, but also a political, aesthetic and ethical dimension, moving from the restricted to the public spaces, and from proactive to reactive dispositions

toward data. Ergo, data literacy can only stem from exploring, interpreting and building understanding around a bigger picture - a complex, integrated perspective on data practices.

## References

- Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society*, 2(2), 205395171559463. <https://doi.org/10.1177/2053951715594634>
- Borgman, C. L., Darch, P. T., Sands, A. E., Pasquetto, I. V., Golshan, M. S., Wallis, J. C., & Traweek, S. (2015). Knowledge infrastructures in science: data, diversity, and digital libraries. *International Journal on Digital Libraries*, 16(3–4), 207–227. <https://doi.org/10.1007/s00799-015-0157-z>
- Carr, W. (2007). Educational research as a practical science. *International Journal of Research & Method in Education*, 30(3), 271–286. <https://doi.org/10.1080/17437270701614774>
- Carretero, S., Vuorikari, R., & Punie, Y. (2017). *The Digital Competence Framework for Citizens With eight proficiency levels and examples of use*. Brussels. <https://doi.org/10.2760/38842>
- Chi, Y., Jeng, W., Acker, A., & Bowler, L. (2018). Affective, behavioral, and cognitive aspects of teen perspectives on personal data in social media: A model of youth data literacy. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 10766 LNCS, pp. 442–452). Springer Verlag. [https://doi.org/10.1007/978-3-319-78105-1\\_49](https://doi.org/10.1007/978-3-319-78105-1_49)
- Corbin, J., & Strauss, A. (2007). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (Vol. 14). Thousand Oaks, CA: SAGE Publications.
- D'Ignazio, C., & Bhargava, R. (2015). Approaches to Building Big Data Literacy. In *Bloomberg Data for Good Exchange* (p. online). Boston, Massachusetts: MIT Press. Retrieved from [https://dam-prod.media.mit.edu/x/2016/10/20/Edu\\_D'Ignazio\\_52.pdf](https://dam-prod.media.mit.edu/x/2016/10/20/Edu_D'Ignazio_52.pdf)
- D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. Boston, Massachusetts: MIT Press. <https://doi.org/10.7551/mitpress/11805.001.0001>
- Eubanks, V. (2018). Algorithms Designed to Fight Poverty Can Actually Make it Worse. *Scientific American*, 319(5), 68–71. <https://doi.org/10.1038/scientificamerican1118-68>
- Freire, P. (2000). *Pedagogy of the Oppressed: 30th Anniversary Edition*. Bloomsbury Academic.
- Gadamer, H.-G. (2013). *Truth and Method*. London: Bloomsbury.
- Gould, R. (2017). Data Literacy is Statistical Literacy. *Statistics Education Research Journal*, 16(1), 22–25. Retrieved from [https://iase-web.org/documents/SERJ/SERJ16\(1\)\\_Gould.pdf](https://iase-web.org/documents/SERJ/SERJ16(1)_Gould.pdf)
- Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, 5(2), 205395171878631. <https://doi.org/10.1177/2053951718786316>
- Hobbs, R. (2020). Propaganda in an Age of Algorithmic Personalization: Expanding Literacy Research and Practice. *Reading Research Quarterly*, 55(3), 521–533. <https://doi.org/10.1002/rrq.301>
- Lehtiniemi, T., & Ruckenstein, M. (2019). The social imaginaries of data activism. *Big Data & Society*, 6(1), 205395171882114. <https://doi.org/10.1177/2053951718821146>
- Lupton, D., & Williamson, B. (2017). The datafied child: The dataveillance of children and implications for their rights. *New Media & Society*, 19(5), 780–794. <https://doi.org/10.1177/1461444816686328>
- Malik, M. M. (2020). A Hierarchy of Limitations in Machine Learning. Retrieved from

<http://arxiv.org/abs/2002.05193>

- Manca, A., Atenas, J., Ciociola, C., & Nascimbeni, F. (2017). Critical pedagogy and open data for educating towards social cohesion. *Italian Journal of Educational Technology*, 25(1), 111–115. <https://doi.org/10.17471/2499-4324/862>
- Mandinach, E. B., & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, 60, 366–376. <https://doi.org/10.1016/j.tate.2016.07.011>
- Markham, A. N. (2018). Critical Pedagogy as a Response to Datafication. *Qualitative Inquiry*, 107780041880947. <https://doi.org/10.1177/1077800418809470>
- Maybee, C., & Zilinski, L. (2015). Data informed learning: A next phase data literacy framework for higher education. *Proceedings of the Association for Information Science and Technology*, 52(1), 1–4. <https://doi.org/10.1002/pra2.2015.1450520100108>
- Milan, S., & van der Velden, L. (2016, October 10). The Alternative Epistemologies of Data Activism. Retrieved from [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2850470](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2850470)
- O’Neil, C. (2016). *Weapons of math destruction : how big data increases inequality and threatens democracy*. New York: Penguin.
- Pangrazio, L. (2016). Reconceptualising critical digital literacy. *Discourse*, 37(2), 163–174. <https://doi.org/10.1080/01596306.2014.942836>
- Pangrazio, L., & Selwyn, N. (2019). ‘Personal data literacies’: A critical literacies approach to enhancing understandings of personal digital data. *New Media and Society*, 21(2), 419–437. <https://doi.org/10.1177/1461444818799523>
- Pangrazio, L., & Selwyn, N. (2020). Towards a school-based ‘critical data education.’ *Pedagogy, Culture and Society*. <https://doi.org/10.1080/14681366.2020.1747527>
- Papavlasopoulou, S., Giannakos, M. N., & Jaccheri, L. (2017). Empirical studies on the Maker Movement, a promising approach to learning: A literature review. *Entertainment Computing*. <https://doi.org/10.1016/j.entcom.2016.09.002>
- Perrotta, C., Gulson, K. N., Williamson, B., & Witzenberger, K. (2020). Automation, APIs and the distributed labour of platform pedagogies in Google Classroom. *Critical Studies in Education*, 00, 1–17. <https://doi.org/10.1080/17508487.2020.1855597>
- Popat, S., & Starkey, L. (2019). Learning to code or coding to learn? A systematic review. *Computers and Education*. <https://doi.org/10.1016/j.compedu.2018.10.005>
- Pöttsch, H. (2019). Critical digital literacy: Technology in education beyond issues of user competence and labour-market qualifications. *TripleC*, 17(2), 221–240. <https://doi.org/10.31269/triplec.v17i2.1093>
- Prinsloo, P. (2020). Data frontiers and frontiers of power in (higher) education: a view of/from the Global South. *Teaching in Higher Education*, 25(4), 366–383. <https://doi.org/10.1080/13562517.2020.1723537>
- Pybus, J., Coté, M., & Blanke, T. (2015). Hacking the social life of Big Data. *Big Data & Society*, 2(2), 205395171561664. <https://doi.org/10.1177/2053951715616649>
- Raffaghelli, J. E. (2018). Educators’ Data Literacy Supporting critical perspectives in the context of a “datafied” education. In M. Ranieri, L. Menichetti, & M. Kashny-Borges (Eds.), *Teacher education & training on ict between Europe and Latin America* (pp. 91–109). Roma: Aracné. <https://doi.org/10.4399/97888255210238>
- Raffaghelli, J. E. (2020a). «Datificación» y Educación Superior: Hacia la construcción de un marco para la

alfabetización en datos del profesorado universitario. *Revista Interamericana de Investigación, Educación y Pedagogía, RIIEP*, 13(1), 177–205. Retrieved from <https://revistas.usantotomas.edu.co/index.php/riiep/article/view/5466>

- Raffaghelli, J. E. (2020b). Is Data Literacy a Catalyst of Social Justice? A Response from Nine Data Literacy Initiatives in Higher Education. *Education Sciences*, 10(9), 233. <https://doi.org/10.3390/educsci10090233>
- Raffaghelli, J. E., & Manca, S. (2019). Is there a social life in open data? The case of open data practices in educational technology research. *Publications*, 7(1). <https://doi.org/10.3390/publications7010009>
- Raffaghelli, J. E., Manca, S., Stewart, B., Prinsloo, P., & Sangrà, A. (2020). Supporting the development of critical data literacies in higher education: building blocks for fair data cultures in society. *International Journal of Educational Technologies in Higher Education*, 17(58). <https://doi.org/https://doi.org/10.1186/s41239-020-00235-w>
- Raffaghelli, J. E., & Stewart, B. (2020). Centering complexity in ‘educators’ data literacy’ to support future practices in faculty development: a systematic review of the literature. *Teaching in Higher Education*, 25(4), 435–455. <https://doi.org/10.1080/13562517.2019.1696301>
- Ricaurte, P. (2019). Data Epistemologies, The Coloniality of Power, and Resistance. *Television & New Media*, 20(4), 350–365. <https://doi.org/10.1177/1527476419831640>
- Sancho-Gil, J. M., Rivera-Vargas, P., & Miño-Puigcercós, R. (2020). Moving beyond the predictable failure of Ed-Tech initiatives. *Learning, Media and Technology*, 45(1), 61–75. <https://doi.org/10.1080/17439884.2019.1666873>
- Stephen-Davidowitz, S. (2017). *Everybody Lies. Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are*. New York: Harper Collins.
- Stewart, B. (2020). The Open Page Project. *Journal of Teaching and Learning*, 14(1), 59–70. <https://doi.org/10.22329/jtl.v14i1.6265>
- Tour, E., Creely, E., & Waterhouse, P. (2021). “It’s a Black Hole . . .”: Exploring Teachers’ Narratives and Practices for Digital Literacies in the Adult EAL Context. *Adult Education Quarterly*, 074171362199151. <https://doi.org/10.1177/0741713621991516>
- Tygel, A. F., & Kirsch, R. (2016). Contributions of Paulo Freire for a Critical Data Literacy: a Popular Education Approach. *The Journal of Community Informatics*, 12(3). Retrieved from <http://www.ci-journal.net/index.php/ciej/article/view/1296>
- van der Zee, T., & Reich, J. (2018). Open Education Science. *AERA Open*, 4(3), 233285841878746. <https://doi.org/10.1177/2332858418787466>
- Wasson, B., Hansen, C., & Nettelund, G. (2016). Data Literacy and Use for Learning when using Learning Analytics for Learners. In S. Bull, B. M. Ginon, J. Kay, M. D. Kickmeier-Rust, & M. D. Johnson (Eds.), *Learning Analytics for Learners, 2016 workshops at LAK* (pp. 38–41). Edimburg: CEUR. Retrieved from <http://ceur-ws.org/Vol-1596/paper6.pdf>
- Wilkerson, M. H., & Laina, V. (2017). Youth Reasoning With Interactive Data Visualizations. In *Proceedings of the 2017 Conference on Interaction Design and Children - IDC '17* (pp. 411–416). New York, New York, USA: ACM Press. <https://doi.org/10.1145/3078072.3084302>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>
- Williamson, B., Eynon, R., & Potter, J. (2020, April 2). Pandemic politics, pedagogies and practices: digital technologies and distance education during the coronavirus emergency. *Learning, Media and*

*Technology*. Routledge. <https://doi.org/10.1080/17439884.2020.1761641>

Willis, J. E., Slade, S., & Prinsloo, P. (2016). Ethical oversight of student data in learning analytics: a typology derived from a cross-continental, cross-institutional perspective. *Educational Technology Research and Development*, 64(5), 881–901. <https://doi.org/10.1007/s11423-016-9463-4>

