Datafixation of education

Pekka Mertala

Abstract
This article problematizes the datafication of education through the concept of datafixation —a wordplay that combines datafication and fixation. The article showcases how the datafication of education presents compelling arguments about the importance and usefulness of data, while remaining silent about the negative consequences of reducing students, learning, teaching, and education to superficial indicators. Such a one-sided view meets the criteria of fixation as an unhealthy attachment. I will also demonstrate that datafixation is not limited to any particular agents, technologies, or educational contexts but, like datafication, is a phenomenon that encompasses various fields and permeates different societal layers from global educational policies to classroom activities. Thus, throughout the article, I will provide examples from various sources, including edtech companies' materials, policy documents, and scholarly articles, to illustrate that they share similar views and rhetoric. Diversity is also evident in the types of data technologies —such as facial recognition technology, learning management systems, and national continuing education platforms— that this paper concretizes datafixation with.

Yes, the title is (or attempts to be) a witty wordplay that combines the words »datafication« and »fixation«. As the wordplay suggests, the argument I am about to make is that the datafication of education meets the criteria of fixation —an obsessive or unhealthy attachment— as defined in the Merriam-Webster Dictionary.[1] The disciplinary home of the present paper is critical educational technology (edtech) research, which —among other objectives— is about »observing emerging technologies, questioning the hype surrounding them, and reflecting on their sociopolitical implications« (Macgilchrist, 2021, p. 243). Put differently, in the forthcoming sections, I will showcase how the datafication of education is rich with evocative arguments about the importance and utility of data that remain silent about the negative consequences of reducing students, learning, teaching, and education into superficial indicators.

I will also demonstrate that datafixation is not limited to any particular agents, technologies, or educational contexts but, like datafication, is a phenomenon that

1 https://www.merriam-webster.com/dictionary/fixation (last retrieved 23.06.2023)

covers various fields and pierces different societal layers from global educational policies to classroom activities. Thus, throughout the article, I will provide examples from various sources ranging from edtech companies' materials to policy documents and scholarly articles to illustrate that they share similar views and rhetoric.[2] Diversity is also present within the kinds of data technologies —facial recognition technology, learning management systems, and national continuing education platforms— this paper concretizes datafixation with. In the second section, the focus is on formal education, namely elementary, secondary, and upper-secondary school. In the third section, the gaze is turned towards lifelong and life-wide learning. The first section, in turn, sets the scene by discussing data discourses (the ways data are spoken about in education) on a broad spectrum.

A discursive approach naturally requires an explanation of how the concept of data is understood and approached in this article. I draw on Russell Ackoff's (1989) classical definition of data as »symbols that represent properties of objects, events, and their environments« for two main (terminological) reasons. The first one is Ackoff's decision to use the term »symbols« rather than numbers, »quantified evidence«, or »raw material« (van Dijck, 2014), which are commonly associated with data (e.g., Bowler et al., 2017; van Dijck, 2014; Kitchin, 2014). The Merriam-Webster Dictionary defines a symbol as »something that stands for or suggests something else by reason of a relationship, association, convention, or accidental resemblance[3]«, which neatly pinpoints the indicative nature of data.

Another merit in Ackoff's definition is the use of the term »representation«, which, according to the Cambridge Dictionary[4], is »a sign, picture, model, etc. of something.« Let us use Belgian surrealist Rene Magritte's canonical painting, »The Treachery of Images« (La Trahison des images), as a device to concretize the utility of Ackoff's definition. The picture shows a pipe. Below it, Magritte wrote, »Ceci n'est pas une pipe,« which is French for »This is not a pipe,« to underline that the painting is not a pipe but rather an image (representation) of it.[5] Similarly, data are representations of certain phenomena —and often way less precise than Magritte's painting. However, »this is not learning« is a statement I do not expect to see in any learning analytics marketing materials anytime soon.

---

2 See Mertala (2021) for further discussion about the rhetorical similarities between edtech companies' marketing material and edtech research papers.

3 https://www.merriam-webster.com/dictionary/symbol (last retrieved 23.06.2023)

4 https://dictionary.cambridge.org/dictionary/english/representation (last retrieved 23.06.2023)

5 https://www.renemagritte.org/the-treachery-of-images.jsp (last retrieved 23.06.2023)

Data Discourses

According to Evgeny Morozov (2013, p. 14), the internet as a technical system has little to do with the mythical and all-powerful internet that is discussed in public discourse. Morozov calls the latter ›internet‹ only within quotation marks. By the same logic, there is also data and ›data‹. Data without quotation marks refers to the realistic properties of data; the fact that data are at best only clues and indicators (or symbols to use Ackoff's vocabulary) of the phenomenon it is trying to capture (Selwyn, 2019). ›Data‹, in turn, refers to the discursive production of data as accurate, infallible, and evidential (see Špiranec et al., 2019).

Public educational (and scholarly, to some extent) discussions tend to be more about ›data‹ than data. The unwavering trust towards data is present in the common public and scholarly arguments that education/instruction should be »data-driven« (e.g., Burroughs, 2020; Kurilovas, 2020; Mandinach, 2012). The choice of words is telling: it is not enough that education is informed by data or that data is used in educational decision making. Instead, data must sit in the driver's seat. Let us examine the topic more closely by using the following excerpt from edtech company LSU as an example.

> »One common misconception about data-driven instruction is that it only focuses on teaching methods that lead to higher test scores. In reality, data-driven instruction looks at the whole picture and uses dynamic student data to gauge individual and classroom comprehension, giving teachers insight into specific adjustments they can make to the curricula to improve student understanding. The use of student data to drive instruction also allows teachers to tailor their teaching methods to encourage student achievement.« (LSU, 2020, n.p.)

The excerpt aims to convince the readers through technical and evocative justifications. From a technical point of view, the extract includes two themes common in the pro-datafication ›data‹ discourse: the claim that data is all-encompassing (»looks at the whole picture«; »gauge[s] individual and classroom comprehension«) and the emphasis on the dynamic nature of digital data. To provide some supporting examples, IBM's Watson Element for Educators promises to give teachers »a comprehensive 360-degree view of each student« (IBM, 2018, n.p.). Many edtech companies also highlight the adaptive qualities of their learning software (Yu & Couldry, 2022), which is a result of the dynamic interplay between the software and students' real-time input and accumulated data. When it comes to the evocative qualities, the excerpt claims that data-driven instruction can »encourage student achievement.«

A dictionary definition for encourage is »to inspire with courage, spirit, or hope«.[6] Thus, the choice of words implies that data-driven instruction enables teachers to support students' agency and empowerment with regards to academic achievement instead of mechanically ›tuning up‹ measurable student achievement like the test scores mentioned in the beginning of the excerpt.

That being said, the excerpt fails to mention any concrete technologies or methods to achieve these objectives, and thus it stays at a rather abstract level. The use of facial recognition technology (FRT) offers a more »hands-on« example. According to its proponents, FRT

> »... can help teachers recognize different student emotions in class, measure their levels of interest, frustration, and comprehension, and use this information to adjust their styles accordingly. With FRT, teachers will be able to change their pace of instruction and tailor their classroom instruction to maximize students' involvement and performance.« (Viatech, 2018, n.p.)

These claims contain numerous problems. Facial expressions are poor indicators of experienced emotions to begin with. Moreover, some FRT systems treat people of color differently, interpreting them as having more negative emotions than white faces and registering them as angrier and more contemptuous (Crawford, 2021). It is also important to note that children's faces go through notable changes during their school years, and the »recalcitrance of the face of the child, its vital malleability, makes it a uniquely difficult subject to model and algorithmically capture« (O'Neill et al., 2022, p. 763). As a result, this excerpt is a classic example of how edtech is claimed to bring »significant value to the learner and their educational institution, even when actual practices do not always reflect such imaginaries« (Hansen & Komljenovic, 2023, p. 101).

The problems with the quote are not restricted to technological shortcomings only. The quote's underlying idea of what education and the teacher-student relationship are about do not stand up to critical scrutiny. Let us examine the different motives for recognizing and paying attention to emotions presented in the quote. It is apparent that the emotion and the factors that led to it have no intrinsic importance, nor is FRT intended to support students in practicing emotion management and self-regulation. Instead, emotions are reduced to factors that support or challenge (academic) learning, which need to be either amplified or deactivated to maximize the effectiveness of learning. The approach is mechanistic and does not meet

6 https://www.merriamwebster.com/dictionary/encourage?utm_campaign=sd&utm_medium=serp&utm_source=jsonld; On the other hand, encourage also refers to one party's »attempt to persuade« another one –which contains a (most likely unintentional) connotation with data-based profiling and steering of peoples commercial and political choices.

the needs of students. If the FRT detects that students are sad or afraid, changing »the pace of instruction and tailor their classroom instruction to maximize students' involvement and performance[7]« (Viatech, 2018, n.p.) is hardly a pedagogically sound response to the situation. Nor is »adjusting their styles accordingly.« To conclude, if we allow data to sit behind the wheel, we may end up finding ourselves in rather unpleasant places and situations —a topic that is discussed more profoundly in the following section.

Ubiquity of Datafication: School as a Data Factory

The history of data production and collection in educational settings is probably as long as the history of formal education: for ages, attendance lists and exam results have been collected in order to track students', teachers', and schools' performance. Additionally, the view of school as a ›data factory‹ is not a fully contemporary phenomenon, but utopias of fully automated schools date back to Simon Ramo's writings from the 1950s (Watters, 2021).

That said, the intertwinement of technological development and competence-oriented and measurement-fixated educational policy (see Mertala, 2020) has increased the intensity of data collection in schools exponentially. We have already touched upon the use of FRT, but learning analytics (Selwyn, 2019) and the use of wearables like heart rate monitors in physical education (Williamson, 2016) are probably more familiar examples. Regardless of the type of technology, data are typically extracted, analyzed, and reported to the teacher/administration automatically. Thus, the student becomes a subject of continuous screening and surveillance —a phenomenon often referred to as dataveillance (e.g., Yu & Couldry, 2022).

Take learning analytics, for example. The regime of constant testing and measurement is —metaphorically and literally— coded into learning analytics' functional logic. The utility and importance of ViLLE learning analytics platform used in about half of Finnish schools (University of Turku, 2020) is justified via an argument that the »automatic analytics enables real-time viewing of information [for teachers and administration] without a separate test.« (Centre for learning analytics, 2019, n.p.). In other words, students unconsciously take a test every time they use the app.

Furthermore, schools' data technologies are often interlinked and interoperable, which means that they communicate and share data with each other (Cone, 2022). This development is called platformization, in which previously distinct functionalities are bundled together (Kerssens & van Dijck, 2021). The most significant and

7 It is worth noticing that both examples use the verb »tailor« to describe teachers' actions.

defining platform is typically the learning management system (LMS), which is used for a whole host of activities, including delivering content, circulating classroom assignments, interacting with colleagues and students, keeping records, making reports, and so on (Pangrazio et al., 2022). Edtech company Knewton, for instance, emphasizes the importance of platform-like LMS by stating that:

> »Today, students walk into classrooms each September as if they were just born. Teachers must learn everything about them from scratch. Knewton-powered apps change this, allowing each student to start courses ›warm‹ by connecting his or her learning history to every app.« (Yu & Couldry, 2022, p.135)

The advertised ability to connect each student's learning history to »every app« is an explicit reference to the previously mentioned idea of interoperability: the LMS serves as a nexus for various applications to share and retrieve data. Furthermore, the excerpt also strongly emphasizes the importance of accumulative data collection (and sharing) throughout students' school-era. Then again, why restrict dataveillance to the sphere of formal education only. It is argued learning analytics applications should be allowed to collect personal data, including online behavior outside learning platforms, as »such data has great potential for understanding and optimizing learning processes« (Ifenthaler & Schumacher, 2016, 933). A Finnish edtech company, Wisma, whose Wilma LMS has two million users in a country with a population of under six million, goes even further and envisions »an LMS on steroids«, which merges not only educational data but also various forms of registry data.

> »Already, it would be possible to collect information on learners not only on their schooling but also on their family, health, and social support, for example. By systematically using this data, predictive and guiding analyses could be made of what is likely to happen next in the life of a particular type of learner and how to respond to it at what stage.« (Salkolahti, 2022, n.d.)

Anyone familiar with the UK A-level exam grading fiasco from 2020 will probably be alarmed when visions of »predictive analyses« of »what is likely to happen in the life of a particular type of learner« are introduced. In the absence of actual exams, Ofqual (The Office of Qualifications and Examinations Regulation) decided to estimate A-level grades by an algorithm using three (data) inputs: 1) The historical grade distribution of schools from the three previous years; 2) The rank of each student within her own school for a particular subject, based on a teacher's evaluation of their likely A-level grade; and 3) The previous exam results for a student per subject (Kolkman, 2020). The model ended up favoring students from independent (fee-paying) schools (Kolkman, 2020), and disadvantaged students were the worst affected as the algorithm copied the inequalities that exist in the UK's education system (Shead, 2020).

That being said, the examples above provide a somewhat one-sided view. While it is true that algorithmic bias and predictive models can (and often do) hold up and even worsen societal inequalities, people are not mere passive »sources« of data but can also use data to create the kinds of self-representations they desire. In the following section these themes are elaborated from the viewpoint of lifelong and life-wide learning.

Data as Representation of Being a Lifelong and Life-wide Learner

»A good worker is an agile learner« (Otala, 2018, 22) —a statement that neatly compresses the master narrative of contemporary work life. Jobs are constantly changing, and people need to complement their previous educational experience through continued learning in different contexts (Desjardins, 2003) to meet new demands. To succeed in the labor market, individuals must build a proactive, self-reliant, and renewing »aura« of lifelong and life-wide learning around them. In other words, one has to show that they can learn throughout the life course (lifelong learning) in formal, informal, and nonformal contexts (life-wide learning). This »showing« often takes place via digital self-representation, a common topic in impression management literature. Job seekers are instructed on how to present the most favorable image to potential employers (e.g., Fertik & Thompson, 2015; Krings et al., 2021; Paliszkiewicz & Mardra-Sawica, 2016) or how to create symbols that represent desired properties of the job-seeking subject, to loosely paraphrase Ackoff's (1989) words. One practical example of this is search engine optimization (SEO):

> »Boosting the SEO of your blog, website, and social media presence will allow recruiters and candidate-searching employers to more easily and quickly discover you; and while of course, you should also be applying to jobs and networking within your industry, better SEO means a higher chance of naturally attracting interested visitors.« (Simmons, n.d.)

The excerpt contains explicit and implicit references to the three modes of self-representation in digital media: written, visual, and quantitative (Walker-Rettberg, 2017). A blog is an obvious reference to written self-representation, and one can write commentaries about timely and topical professional themes to represent oneself as an expert who keeps up-to-date via informal and nonformal learning. Visual self-representation, in turn, is about using attractive images to increase interest. Using a profile photo with a younger appearance can increase job offers for older job seekers (Krings et al., 2021).

Lastly, quantitative self-representation is touched upon by mentioning networking. Research suggests that employers see a wide network of contacts on LinkedIn (an online service that profiles itself as a space for professional networking and showcasing skills) as a positive and desirable feature (Zide et al., 2014), and the reputational literature also recommends joining different groups (Paliszkiewicz & Mardra-Sawica, 2016). Following Anna Sfard's (1998) infamous metaphor of learning as participation, a wide professional network symbolizes possibilities for multifocal interaction, which enables boundary crossing and »out-of-the-box« thinking. Of course, a thousand contacts on LinkedIn do not tell us anything about the density or interactivity of the network. It says that 1,000 people have responded positively to a person's invitation to network. And why shouldn't they because a positive response also increases their own network.

Intentional digital self-representation can also take place in the realm of formal education. In Finland, all higher education institutions are involved in the Digivisio 2030 project, which aims to create an internationally esteemed ecosystem for lifelong learning. Part of the ecosystem is a service called »My Data«, which, based on the somewhat obscure public description, functions as a type of learning/competence portfolio (Digivisio, 2030, n.d.). Users can use the »My Data« service to represent themselves as lifelong learners by logging into the system, regularly browsing the course tray, and naturally, attending courses to produce data points that represent them as active seekers of in-service training.

The above scenarios are partly hypothetical in nature. However, given that various stages of the recruitment process have become more automated, it is important to ponder these scenarios critically. Currently, many companies offer services where AI mines data from platforms such as LinkedIn, Facebook, Instagram, Twitter, and others to identify potential active and passive employee candidates that match the defined needs of the company purchasing the service (Black & van Esch, 2020). The open job opportunity becomes known only to the candidates identified through the extraction process. In this case, a potential employee will only come under the radar of the company if their digital self-representations resonate sufficiently with the search parameters. Being the best candidate is less important than looking like the best candidate.

Concluding remarks

In the title of this article, I described the nature of datafication as datafixation —an obsessive trust in the accuracy and usefulness of data, which at worst has unhealthy consequences for education and people, in general. I am aware that a reader who disagrees with my arguments may interpret this essay as an expression of obsessive concern, which is another form of the dictionary definition of fixation.[8] To avoid unnecessary confrontation (and one-eyed criticism that offers no alternative imaginaries, see Morozov, 2015), I stress that a critical stance does not mean that I do not see data technologies as having something to offer schools and education. For example, applications using location data have been found to increase mobility during the school day —at least during the period of the study (Koivisto et al., 2020). However, sustainable benefits can only be achieved by recognizing and addressing the limitations, blind spots, and biases of data technologies and the (material and social) consequences these issues cause.

Additionally, it is equally important to pay critical attention to the ways in which data is spoken about. Metaphors like »data mining«, data is »raw material«, and »new oil« juxtapose data with natural resources that exist regardless of the people (see also van Dijck, 2014). That, however, is not the case with educational data. There is no data for learning analytics to analyze if the student does not log in and click the mouse, tap the screen, or press a key. The same principle applies to juxtaposing data as direct measurement (see Selwyn, 2019; Špiranec et al., 2019). The data from students' interactions with learning analytics is not undisputed evidence about learning (either as a process or an outcome). Instead, they are (mere) representations of indicators of learning, based on the level of correspondence between »digital behavior objects« (students' actions) and »digital content objects« (platform's expectations) (see Hansen & Komljenovic, 2023). Furthermore, as I have exemplified, on some occasions people can also game the system by creating data points that symbolize a desired representation, which may have little to no correspondence with the real world. Indeed, Magritte's message about the treachery of images is something that all people engaged with data and education should be aware of and open about.

As a related notion, opportunities for critical discussions and changemaking should not be restricted to scholarly and/or (macro) political level agents only. It is crucial to make the limitations, blind spots, and biases of datafication visible to students as well and engage them in seeking solutions to overcome these issues. In the end, students are the primary group to be affected by datafication (and datafixation) of education, and they have the right to have a word to say. One concrete way to pursue this objective would be data literacy education that puts schools (and other educational contexts) data practices and technologies under similar critical investigation as, say, the recommendation algorithms of Google's services, which are a common case of contemporary data/media literacy education (see Grosman et al., 2022). Given that Google dominates also the edtech market (Schoolov, 2019), it would be a well-justified case for critical exploration of educational datafication as well.

8 https://www.merriam-webster.com/dictionary/fixation (last retrieved 23.06.2023)

References

Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, *16*(1), 3–9.

Black, J. S., & van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? *Business Horizons*, *63*(2), 215–226.

Bowler, L., Acker, A., Jeng, W., & Chi, Y. (2017). »It lives all around us«: Aspects of data literacy in teen's lives. *Proceedings of the Association for Information Science and Technology*, *54*(1), 27–35.

Burroughs, A. (2020). *Why K–12 Schools Should Establish a Data–Driven Culture. EdTech: Focus on K–12.* https://edtechmagazine.com/k12/article/2020/04/why–k–12–schools–should–establish–data–driven–culture–perfcon (last retrieved 23.06.2023)

Centre for Learning Analytics (2019). *Example 4: Automating learning analytics*. https://en.learninganalytics.fi/analytics#case__4 (last retrieved 23.06.2023)

Cone, L. (2023). The platform classroom: Troubling student configurations in a Danish primary school. *Learning, Media and Technology*, *48*(1), 52–64.

Crawford, K. (2021). *The atlas of AI*. Yale University Press.

Desjardins, R. (2003). Determinants of literacy proficiency: A lifelong-lifewide learning perspective. *International Journal of Educational Research, 39*(3), 205–245.

Digivisio 2020 (n.d.). *New services developed in Digivisio 2030*. https://digivisio2030.fi/en/new–services–developed–in–digivisio–2030/ (last retrieved 23.06.2023)

van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197–208.

Fertik, M., & Thompson, D. (2015). *The reputation economy: How to optimise your digital footprint in a world where your reputation is your most valuable asset*. Hachette UK.

Grosman, J., Jacques, J., & Collard, A. S. (2022). »The Beatles with the lower score, it breaks my heart«: Framing a media education response to datafication and algorithmic recommendations in digital media infrastructures. In L. Pangrazio & J. Sefton-Green (Eds.), *Learning to Live with datafication: Educational case studies and initiatives from across the world* (pp. 135–152). Routledge.

Hansen, M., & Komljenovic, J. (2023). Automating learning situations in EdTech: Techno-commercial logic of assetisation. *Postdigital Science and Education, 5*(1), 100–116.

IBM (2018). *IBM Watson Education Classroom helps teachers deliver personalized learning that can improve student outcomes*. https://www.ibm.com/common/ssi/cgi–bin/ssialias?htmlfid=897/ENUS218–010&infotype=AN&subtype=CA (last retrieved 23.06.2023)

Ifenthaler, D. & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development* *64*(5), 923–938.

Jokisaari, O-J. (2004). Elinikäinen oppiminen–häpeä ja menetetty vapaus. *Aikuiskasvatus, 24*(1), 4–16.

Kerssens, N., & Dijck, J. V. (2021). The platformization of primary education in The Netherlands. *Learning, Media and Technology*, *46*(3), 250–263.

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.

Koivisto, K., Kallio, J., Kulmala, J., Hakonen, H., Tammelin, T., & Koski, P. (2020). Mobiilisovelluksen opetuskäytön yhteys kahdeksasluokkalaisten fyysiseen aktiivisuuteen koulupäivän aikana. *Liikunta & Tiede*, *57*(5), 115–122.

Kolkman, D. (2020). »F**k the algorithm«?: What the world can learn from the UK's A–level grading fiasco. https://blogs.lse.ac.uk/impactofsocialsciences/2020/08/26/fk–the–algorithm–what–the–world–can–learn–from–the–uks–a–level–grading–fiasco/ (last retrieved 23.06.2023)

Krings, F., Gioaba, I., Kaufmann, M., Sczesny, S., & Zebrowitz, L. (2021). Older and Younger Job Seekers' Impression Management on LinkedIn. *Journal of Personnel Psychology* *20*(2), 61–74.

Kurilovas, E. (2020). On data–driven decision–making for quality education. *Computers in Human Behavior*, *107*, doi.org/10.1016/j.chb.2018.11.003.

LSU Online & Continuing Education (2020). *How Educators Can Use Student Data to Drive Instruction*. https://online.lsu.edu/newsroom/articles/how–educators–can–use–student–data–drive–instruction/ (last retrieved 23.06.2023)

Macgilchrist, F. (2021). What is »critical« in critical studies of edtech? Three responses. *Learning, Media and Technology*, *46*(3), 243–249.

Mandinach, E. B. (2012). A perfect time for data use: Using data–driven decision making to inform practice. *Educational Psychologist*, *47*(2), 71–85.

Mertala, P. (2020). Data (il)literacy education as a hidden curriculum of the datafication of education. *Journal of Media Literacy Education*, *12*(3), 30–42.

Mertala, P. (2021). »It is important at this point to make clear that this study is not ›anti-iPad‹«: Ed-Tech speak around iPads in educational technology research. *Learning, Media and Technology* *46*(2), 230–242.

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism*. Public Affairs.

O'Neill, C., Selwyn, N., Smith, G., Andrejevic, M. & Gu, X. (2022). The two faces of the child in facial recognition industry discourse: Biometric capture between innocence and recalcitrance. *Information, Communication & Society* *25*(6), 752–767.

Otala, L. (2018). *Ketterä oppiminen. Keino menestyä jatkuvassa. muutoksessa*. Kauppakamari.

Paliszkiewicz, J., & Mądra–Sawicka, M. (2016). Impression management in social media: The example of LinkedIn. *Management*, *11*(3), 203–212.

Pangrazio, L., Stornaiuolo, A., Nichols, T. P., Garcia, A., & Philip, T. M. (2022). Datafication meets platformization: Materializing data processes in teaching and learning. *Harvard Educational Review*, *92*(2), 257–283.

Salkolahti, S. (2022). *Wilmankin dataa voidaan hyödyntää – suuria hyötyjä kasvatustyöhön*. https://www.wilma.fi/ajankohtaista/datan–hyodyntaminen–kasvatustyossa/ (last retrieved 23.06.2023)

Schoolov, K. (2019). *Google is winning in education, but Apple and Microsoft are battling for market share*. https://www.cnbc.com/2019/03/20/apple-google-microsoft-are-battling-for-dominance-in-education.html, (last retrieved 23.06.2023)

Selwyn, N. (2019). What's the problem with learning analytics? *Journal of Learning Analytics*, *6*(3), 11–19.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), 4–13.

Shead, S. (2020). *How a computer algorithm caused a grading crisis in British schools.* https://www.cnbc.com/2020/08/21/computer–algorithm–caused–a–grading–crisis–in–british–schools.html (last retrieved 23.06.2023)

Simmons, J. (n.d.) *Here's how to boost your job search SEO*. https://www.monster.com/career–advice/article/how–to–boost–your–job–search–seo–and–why–you–need–it–to–get–hired (last retrieved 23.06.2023)

Špiranec, S., Kos, D., & George, M. (2019). Searching for critical dimensions in data literacy. *Information Research* 24(4), n.p.

University of Turku (2020). *Use of ViLLE Learning Platform Developed at the University of Turku Increased Manifold Due to Corona Circumstances*. https://www.utu.fi/en/news/news/Use–of–ViLLE–Learning–Platform–Increased–Manifold–Due–to–Corona–Circumstances (last retrieved 23.06.2023)

Viatech (2018). *Maximizing classroom performance with facial recognition Technology.* https://www.viatech.com/en/2018/08/maximizing–classroom–performance–frt/ (last retrieved 23.06.2023)

Walker-Rettberg, J. (2017). Self–representation in social media. In J. Burgess, A. Maverick & T. Poell (Eds.), *SAGE handbook of social media* (pp. 429–443). Sage.

Watters, A. (2021). *Teaching machines: The history of personalized learning*. MIT Press.

Williamson, B. (2016). Coding the biodigital child: the biopolitics and pedagogic strategies of educational data science. *Pedagogy, culture & society*, *24*(3), 401–416.

Yu, J., & Couldry, N. (2022). Education as a domain of natural data extraction: analysing corporate discourse about educational tracking. *Information, Communication & Society*, *25*(1), 127–144.

Zide, J., Elman, B., & Shahani–Denning, C. (2014). LinkedIn and recruitment: How profiles differ across occupations. *Employee Relations*, *36*(5), 583–604.