



Welfare Data Society? Critical Evaluation of the Possibilities of Developing Data Infrastructure Literacy from User Data Workshops to Public Service Media

Jenni Hokka

INTRODUCTION

Datafication has changed society and the economy in fundamental ways, blurring long-established social and institutional divisions (Constantiou & Kallinikos, 2015). The whole of human life is transmuted into data streams and is in danger of being exposed to continuous tracking either by profit-seeking companies (Couldry & Mejias, 2018) or government agencies (Dencik et al., 2016). Datafication allows companies to predict and even modify human behaviour as a means of producing revenue and gaining market control leading to what Shoshana Zuboff refers to as “surveillance capitalism” (Zuboff, 2015). Numerous scholars have raised concerns in regard to how this situation leads to surveillance and violations of the right

J. Hokka (✉)
Aalto University, Espoo, Finland
e-mail: jenni.hokka@aalto.fi

© The Author(s) 2022
A. Hepp et al. (eds.), *New Perspectives in Critical Data Studies*,
Transforming Communications – Studies in Cross-Media Research,
https://doi.org/10.1007/978-3-030-96180-0_12

to privacy and how this may represent a serious threat to democracy and equality (Couldry & Mejias, 2018; van Dijk, 2014; Gangadharan, 2017; Gurumyrthy & Bharthur, 2018; Kennedy & Moss, 2015). It is very difficult, even impossible, for user-citizens to gain full knowledge of the personal data that corporations keep on them. As Zuboff notes, “Surveillance capitalism thrives on the public’s ignorance” (2015: 83). However, in a datafied society in which data-intensive logics and practices have penetrated every aspect of human life (Mayer-Schonberger & Cukier, 2013), people have become accustomed to applications that make everyday life more convenient or even offer ways of earning a living. As Mai suggests, it is now virtually impossible to perform daily activities without giving away personal information which is then capitalised upon by either private enterprises, such as data brokers, or used by public organisations (Mai, 2006). The European Union has tried to protect its citizens by establishing the General Data Protection Regulation (GDPR), which went into effect in 2018. A number of earlier research projects (Selwyn & Pangrazio, 2018; Büchi et al., 2017; Park, 2013) have shown how people, on average, have a very weak understanding of exactly how their personal data is collected, linked, used, sold, and re-sold. Furthermore, as personal data is combined into data packages and sold to different parties by data brokers, it is nearly impossible for even a knowledgeable user to comprehend which parties have access to his or her data. As Micheli et al. (2018) suggest, the ability to protect his or her private data and minimise their ‘digital footprints’ should now be understood as an essential part of digital equality along with digital skills and online access. At present, only people with high levels of computational skills and expertise in data mining have access to data and data analytics tools which means that data power is concentrated within just a few, elite commercial companies such as Google, Facebook, and Amazon (Kennedy & Moss, 2015). Consequently, the GDPR regulation does not help if users do not understand that their data is being tracked and re-sold or how exactly this is done and give their informed consent without knowing what that actually means.

Because of the non-transparent nature of practices such as data mining and user tracking by online applications and platforms, there is, as a number of researchers have suggested, an urgent need to increase the level of digital literacy (Gray et al., 2018; Pybus et al., 2015; Park, 2013). The definition of digital literacy varies a great deal. According to Iordache et al. (2017), the concept of digital literacy most often includes three facets: knowledge, skills, and competence, in which knowledge refers to an

understanding of the available digital tools, skills to the practical capabilities to use them and competence to the ability to use the knowledge and skills in different situations. Digital literacy is now seen as a crucial citizenship capability and most European countries have made digital competence a part of basic education, though, not so often as an independent subject but as a cross-curricular theme. Nearly thirty European education systems also mention data and privacy as one aspect of digital competence. For this reason, data privacy and the protection of personal data have recently become an ever more present and valued part of digital literacy (Iordache et al., 2017: 20). Yet, how data and privacy issues are taught in schools varies a great deal between European countries, districts, and even among individual schools depending both on the interpretations of the concept and on teachers' own digital abilities. The practical examples from the European Commission report on digital competence teaching in European education vary from strong passwords to legal issues in sharing information (European Commission/EACEA/Eurydice, 2019: 41–42).

Alternatively, there have been a number of public pleas from both the public and private sectors to develop not just digital literacy but data literacy. Nowadays, many European curricula also mention data literacy, but in basic education data literacy is understood simply as skills “to analyse, compare and critically evaluate the credibility and reliability of sources of data, information and digital content” (European Commission/EACEA/Eurydice, 2019: 38). According to this simplified definition, data literacy could be conceived as part of digital literacy. Yet, to be precise, data literacy overlaps digital literacy to some extent but comprises a different combination of skills and knowledge. A certain level of digital skills is inherently necessary to improve data literacy. However, while digital literacy emphasises general digital skills that are needed for creating, finding, and analysing digital content by using different kinds of software (Iordache et al., 2017), data literacy refers to the technical skills and statistical and informational literacy needed to produce, use, and interpret computational data (Gray et al., 2018). Data literacy is not a skill that can be learned overnight: it is a complicated knowledge framework that also requires statistical literacy, an understanding of the ethics of using data, and the ability to change the tools used according to purpose or discipline (Wolff et al., 2016). Because of this complexity it does not seem probable that data literacy would become a general skillset in the near future.

It is certainly reasonable to try to improve a population's digital literacy. Yet, digital literacy skills alone do not contribute to people's

understanding of datafication's political, economic, and legislative conditions. As proposed by Gray et al., there is a clear need for data infrastructure literacy that would "not only equip people with data skills and data science but also to cultivate sensibilities for data sociology, data culture, and data politics" (2018: 1). This is a vastly different skill to data literacy, as it requires more political and economic knowledge rather than more technical knowledge such as statistical literacy or coding skills. Data infrastructure literacy is essential in light of data justice (Dencik et al., 2016; Taylor, 2017), as it would help citizens of all ages, genders, and educational backgrounds grasp the full societal effects of datafication and then take a stand on fair conditions in data practices. Only through a more widespread understanding of datafication among the general public do awareness of the significance and vastness of data-gathering in our present societies, political discussion, and democratic decision-making about the conditions and regulation of datafication become possible. In a Nordic welfare society, the state has traditionally played a significant role in promoting equality and democracy among its citizens. By providing universal public services such as low-cost childcare and free education, the Nordic welfare society aims to help people gain skills and abilities that enable them to become full members of society (Holmwood, 2000; Kangas & Kvist, 2013). According to Hänninen et al., the Nordic welfare state has four dimensions: personal autonomy, participation, inclusion, and sustainability. "Autonomy refers to human condition [...] in which she is able to master and manage her own life and decisions. [...] Participation refers to a mode of action which influences people in a common endeavour to change their circumstances. [...] Inclusion refers to a state of circumstances in which all involved are so related that they belong together in such a fashion that they contribute to according to their own capacities. Sustainability refers to complex processes which relate people to each other and balance their relations with the environment helping them face (with precaution) uncertainty and contingency" (2019: 5).

How can welfare state thinking be applied to a datafied society, then? In a welfare data society the citizen should be free to master and manage her/his personal data. S/he should be capable of taking part in decisions on how data gathering practices are organised, regulated, and supervised as they now form one of society's key functions. All citizens' capability to participate in decision-making about the rules of datafication should be ensured through universal education provided by the public sector. Through digital literacy and data infrastructure literacy education, citizens

would be more capable of taking precautions to protect their privacy and control the use of their personal data.

Well before the more intense datafication of everyday life Nordic countries have constituted a special model of the “media welfare state” (Syvertsen et al., 2014). In practice, the media welfare state has meant a policy in which all citizens have been granted universal access to education and information so that there exists equal opportunity for the understanding of the society in which they live. The policy has been successful in the sense that even in the present, platformised media environment, Nordic public service companies, such as NRK (Norway), DR (Denmark), and YLE (Finland) all reach a clear majority (roughly 60–90 per cent depending on the way of counting) of the population on a daily basis (Enerhaug, 2019; DR’s public-service redegørelse, 2019; Nokela et al., 2019). Public service media have been an essential part of the (media) welfare state, as they should encourage participation and the inclusion of all citizens in the political and cultural public spheres (Syvertsen et al., 2014: 7). Furthermore, public service media have been an essential part of cultural policy that has aimed to diminish the influence of global market forces (ibid.: 25–28). In the previous period of mass media, global market forces have mostly referred to international cable channels and production companies. Now, in the present datafied and platformised media environment, the strongest global market forces are undoubtedly the so-called ‘Big Tech’ firms such as Google, Amazon, Facebook, Apple, and Microsoft (collectively *known* under the acronym *GAFAM*).

In this chapter, I claim that European public service media should take on new responsibilities in light of the datafied and platformised society. It should fulfil its mission by educating people of all ages to increase their general levels of digital literacy and data infrastructure literacy, and in this way, empower citizens to take part in determining how datafication could operate equitably. By ensuring that citizens understand the social and political effects of datafication, PSM could enhance citizens’ capability to make informed decisions and protect themselves as users, and more importantly, to form an informed opinion on how data tracking and sharing should be regulated. In the long term, citizens should be able take part in discussing new options that pertain to the present situation in which a user is quite powerless in relation to data mining. In this chapter, I discuss the opportunities for PSM to raise the general level of digital and data infrastructure literacy. Although most other European public service broadcasters (such as ARD/ZDF or France TV) apart from the BBC play a

somewhat minor role in their media market than their Nordic counterparts, they too could adopt a new role in providing adult populations both practical, digital literacy skills and a nuanced understanding of the political, societal, and cultural conditions and outcomes of datafication.

I first present the results from our research workshops organised in cooperation with the Finnish Broadcasting Company, YLE, in which participants were both educated on datafication and interviewed about their thoughts and experiences regarding datafication. The results of the workshops reveal the wealth of challenges faced by digital literacy education. In the second part, I describe what kind of educational content YLE already provides for adult citizens related to datafication and digital literacy. I then discuss whether using public service media to increase awareness of datafication and to develop data infrastructure literacy could be one of the essential large-scale practical solutions needed to tackle the imbalanced power structure between social media platforms and users.

NOTIONS OF DIGITAL LITERACY BASED ON USER DATA WORKSHOPS

The first part of the collaboration with YLE was to organise workshops with ‘average’ users who had no special education related to ‘big data’ such as programming or data-analytics. Methodologically, the research followed an emancipatory and educational action research approach (Carr & Kemmis, 2009). The workshops had several, overlapping features. First, the aim was to discuss with the participants their worries, thoughts, and hopes about datafication, especially in regard to the use of their personal data. Second, we wanted to examine how well participants protected their privacy and how willing they were or capable of doing that. In this way, we wanted to increase qualitative understanding, building on previous research (Büchi et al., 2017; Park, 2013; Kennedy et al., 2017; Selwyn & Pangrazio, 2018; Ruckenstein & Granroth, 2019), on users’ perceptions of datafication and their digital capabilities. Third, we provided education on data collection practices and instructions on how to protect privacy online if participants were interested in learning those skills. Fourth, YLE examined what topics and perspectives of datafication interested different audiences. This had the further intention of producing educational media content about datafication to develop digital literacy, and I will discuss this aspect further in the second section. Fifth, we wanted to raise users’

general awareness of datafication and data collection practices to increase their level of data infrastructure literacy. Sixth, we wanted to discuss and develop with the participants the possibilities of alternative data regimes and practices by offering them three alternative visions.

There was a total of six workshops which were partly organised in cooperation with the YLE Creative Content Unit and, more specifically, the Head of Development at YLE, Raimo Lång. The first two workshops were more experimental and concentrated more on acquiring knowledge on users' interest in YLE's potential datafication content, but they also included some of the same questions as the last four workshops. In the last four workshops, which were led by our researcher group and constitute the main research material in this section, we had an identical pattern of action. Each of the four workshops had four to seven participants, both female and male, adding to a total of twenty-five participants. Most were young adults of varying educational backgrounds, but one workshop was arranged for people in their seventies. Even though the participants differed in gender, age, and education, their answers and reactions exhibited similar patterns.

During these workshops, users were asked to familiarise themselves step by step with their Google account's privacy settings and the data that Google had collected on them. In addition, the participants were asked to try the Disconnect application, which informs users on how many third parties were gathering data on them through the websites they visited. After each section, the participants were asked to answer related questions on a Google Forms questionnaire. The reason for asking them to answer in a literal form first was to diminish the effect of participants' views on each other. After participants had sent their answers online to us after each section, they were also asked to discuss their answers with us and other participants. The questions concerned their thoughts and feelings regarding the data that Google collected on them and if they now wanted (or did not want) to change their privacy settings and why. There were also questions about their thoughts on the results of the Lightbeam and Disconnect applications that show the number of third-party requests on each site. In the end, the participants were given a more complex question on their vision for data collection practices of platforms, online applications, and websites in the future.

At the beginning of each workshop, we asked participants if they were slightly, very, or not at all concerned about the gathering of private data on the web. Most people were slightly concerned, except for the group of

women in their seventies, who were in the main very concerned. During the workshops in which participants were introduced to how their online behaviour was tracked, the participants described a range of feelings, including contradictory ones. Nearly half the participants (eleven) described feeling frightened, concerned, confused, shocked, startled, and even angry upon learning the amount of data that Google or third parties were collecting online. In particular, the data from Google Maps, their Google browsing history and/or the results of the Lightbeam and Disconnect applications seemed to be unpleasant for many participants.

After looking at her browsing history in Google's My Activity section, one participant commented:

I am slightly frightened, as all the search words that I have used tell something about me and my life situation, and I am concerned where this data can be further transmitted.

A participant in her twenties expressed her concerns on Google Maps:

It is awful how well Google knows where I have been. Somebody could easily follow my movements through Google. I started feeling insecure.

Another participant in her twenties commented on her Disconnect results:

Confusingly, many parties follow my every step on the web. There are some parties that, luckily, I am able to prevent from gathering data, but there are way too many parties that I can't prevent from doing that. I would like for the Internet to be the anonymous world that people still describe it as being. This feels like a rough coming back to reality.

A participant in his thirties responded to his Disconnect results by saying:

This is very confusing! It was not surprising that advertising or data analysis companies were tracking data on popular web sites, but I was really confused to notice that Imgur was in contact with a Russian news site. What should I think about this?

In two answers, the participants stated that they were not concerned on behalf of themselves and their own data but saw the vastness of

data-gathering practices as worrying or interesting on a more general, sociopolitical level, especially regarding attempts to influence and/or interfere with elections. One respondent stated a feeling that “there is a risk that at some point, delivering private data may go too far” but felt that the limit had not yet been surpassed. One participant stated that she would have been devastated by this kind of data collection if she had been asked before the present situation, but she was now used to it. In addition, ten respondents had already changed their My Activity settings in some respect to protect their privacy. Two participants stated that they used the incognito setting when browsing, and one had already installed an ad-blocking application.

However, concern and upset were not the only feelings people experienced when looking at their Google data. Nine respondents felt indifferent about at least part of their results, stating that they did not consider this kind of information dangerous, or alternatively, that they were already aware of these data collection practices. It was mainly browsing and location history that these respondents felt were safe to give to Google. Yet it is noteworthy that feelings related to privacy could vary according to the type of information. For example, one participant did not consider location history to be dangerous, but was startled to find that Google had a recording of her voice and that Google was still exchanging data with some third-party applications, such as games, that she had removed from her mobile phone long before: “Why does a mobile game have access to my Google Drive?”

A few respondents underlined the convenience to the user that data collection practices enabled and wanted to continue providing this data in the future as well. Google Maps was seen as especially helpful, for example, when jogging or driving. Many participants considered the location history aspect of Google Maps beneficial because they could see which places or restaurants they had been to, and for many of them the location history acted as a kind of personal diary that stirred pleasant memories. As Mark Andrejevic already stated in 2011, Google has come to be treated much like a public service both by users and institutions. The participants’ responses demonstrate how Google Maps especially, along with Google Search, has become a necessary and unquestionable part of everyday life.

The Ads personalisation section in particular generated mixed feelings from the respondents. The results from our workshops are in line with a previous study by Ruckenstein and Granroth (2019). Similar to their interviewees, targeted advertising was, in the first place, the main or even

the only feature from which people could observe that their actions online were being tracked. In our workshops, a number of respondents were amused by their list of interests for Ads personalisation either because of its accuracy or non-accuracy, but it could cause strong negative emotions too. Over half (fourteen) of the respondents, even those participants who were concerned about their privacy, wanted to have personalised rather than non-personalised advertising. When seeing their lists, many revised them to better correspond with their interests. Participants were, therefore, voluntarily providing more data to Google and in a more targeted way. It appeared that the users responded emotionally to their list of (commercial) interests as markers of their identity and, because of that, felt a need to make it correspond to their ‘real’ interests. One participant was very irritated when seeing her presumed list of interests because she had thought that she had changed her settings to prevent personalisation but also because her profile was “generic and erroneous”.

Several times during the workshops, participants expressed their surprise about their own Google settings that they thought they had set otherwise or for which they did not recall taking any action. Several respondents were also surprised that some third-party applications they had used still had access to their Google account, and they either did not remember or had not realised they were giving this permission when using their Google account for some third-party applications. In everyday use, people easily forget or are not capable of protecting their privacy even if they would consider that important on a more general level. As Park et al. (2018) have noted, privacy regulations are based on the assumption of a rational user—but for the most part, people do not actively and rationally weigh the consequences of their every step online from a privacy point of view, especially when the online environment is structured to provide all kinds of pleasurable feelings from sharing one’s personal data.

Furthermore, especially in the group in which the participants were in their seventies, the practical skills to protect one’s privacy were quite low. For a few, basic skills such as using several different browser windows at the same time were difficult, and some had not realised how a simple online toggle button works—even though they actively used many kinds of online applications and services. Many of them had difficulties managing junk mail, and they asked us, the organisers, what they should do to block it. To be able to protect one’s privacy, therefore, a user should have fairly good digital skills (Büchi et al., 2017). The participants in this group felt great unease with these practical problems the logic of which they did

not understand. As a previous study shows (Schreus et al., 2017), the socio-emotional aspects and the level of self-efficacy are crucial factors to bear in mind when thinking of digital literacy education among older users. This is challenging but still very important, as, especially when thinking of users with low digital skills, the idea of users' 'informed consent' seems very unrealistic.

In general, participants were suspicious about the data-gathering practices of social media platforms such as Facebook or Instagram but had little knowledge of the data collection practices carried out by third parties on ordinary websites. This is probably related to the fact that news media have reported many of the privacy scandals related to social media applications; a few participants even mentioned that they had become more cautious after the scandal related to Facebook and Cambridge Analytica. However, the news media, which itself takes part in data collection through its own sites and applications (Helberger, 2016; Turow, 2011; Ruohonen & Leppänen, 2017), have not been so eager to inform people of the regular practices of data-selling to third parties that they also use.

Even for those who had adjusted their Google settings in efforts to protect their privacy in some ways, the vast nature of data collection by third parties through ordinary websites came as a surprise. This particularly held true for the youngest and the oldest participants. Even in the group of the most educated participants (who had all completed their master's studies at university), many had no prior knowledge of how much personal data Google had about them and had not checked their My Activity information before. Research from the last decade shows that, on a general level, most users do not fully comprehend how cookies work (Ha et al., 2006; Jensen et al., 2005), and users grossly overestimate their knowledge of cookies on a research survey (Jensen et al., 2005). In addition, many people tend to think that if one wants to use a certain site, one has no other option than to accept all the cookies (Selwyn & Pangrazio, 2018). Our preliminary findings based on workshops suggest that even though people are capable of linking privacy notices and cookies, they mostly do not realise that cookies do not only share data with the provider of the site but also with a number of other data analysis and advertising companies. As noted (Luzak, 2014), there is no point in asking users for their 'informed consent' to share their data if a majority of users are not aware of cookies or do not understand how they work.

At the end of the two-hour workshops, the participants were introduced to three options for organising their online world if the internet

were invented now and there were no personal data already shared through applications and services. The reason for setting the imaginary frame was to create free space for the participants to consider an ideal scenario without cynically figuring out how much of their data is already available to outside interests. The idea was that this kind of ideal scenario would offer guidelines for future discussions on how to make the present situation better, both for present and future user generations. The three options were:

1. As the user, I share my data so that I can use the applications and sites I want. My personal data can also be shared with third parties. The service provider may come from any country, and it acts upon its information security legislation regarding my data.
2. Service providers have permission to sell my personal data to third parties, but I am able to see what kind of data each service provider and company have about me through my personal *data bank*. Through my data bank, I am able to remove my data from these companies and service providers when I stop using their services or applications.
3. All the online services and applications work through subscriptions. I pay a monthly fee for each service and application, and my data is not sold to third parties. However, I understand that this would make the innovation of new applications and services more difficult.

The first option responds to the situation before the GDPR. The second option was developed by the research group from the ideas of the My Data movement (Lehtiniemi, 2017; Lehtiniemi & Ruckenstein, 2019). The third offers a realistic option in which no data would be exchanged for services and applications, but these would be financed through user payments.

With the exception of four respondents, every participant chose the second option. The second option was often considered as a reasonable compromise. Most participants commented that they still wanted to use free services, but at the same time they wanted to know about the use of their data. Many comments underlined that users should have the right to control the use of their data:

I think that the user whose data forms part of service providers' increase in value, should have a right to know what data is collected and to have a right to control its use. Overall, it is important that data collection is done according to lawful principles. Preventing the irresponsible data gathering by ser-

vice providers should not only be users' responsibility. Even though data collection would be performed according to the rules, the user should have the right to control their data.

This option would secure that service providers would compete with each other to offer better services surrounding shared data. So, users could directly influence service providers.

Even those participants who, in general, did not consider sharing their personal data to be harmful stated that they felt uneasy about the option that service providers would act according to the information security legislation of the provider's origin. Apparently, many of these participants had not realised that this was the situation before the GDPR.

Younger participants in particular noted that they would not have enough money to pay for every application, and a few respondents also thought that subscribing and paying for every application would be very inconvenient. Two participants said that they did not choose the third option but the second because they regarded it as important that there still exists the possibility of developing new applications through data collection; one mentioned that she would happily share her data to be used by health companies. Projects in which organisations or companies have donated their data for the public good have been implemented (Susha et al., 2019; Petersen, 2019; Taylor, 2016), and the participant might have been aware of the idea, as in Finland The Finnish Institute for Health and Welfare already openly shares their anonymous statistical health data concerning, for example, the number of visits and different procedures in each county (THL Open Data). However, when private companies seeking profit develop innovations from open data, the definition of 'public good' might become tenuous, and again, it is questionable as to whether or not most users really understand how revealing their data might be when giving their consent (Taylor, 2016; Lindman & Kuk, 2015).

Yet, some respondents also criticised the second option, even though they had chosen it. Since they thought the data bank option would also be very risky, they offered a new, more developed version of this option:

I would choose the second option, if I could choose a data bank that agrees with third parties that they would have access to my personal data but would not own it. This way data could really be removed so that I would vanish and cease to exist from everybody. I would also like to limit in advance the selling of my data to third parties. I would also like to have the possibility of

making several agreements with different data banks so that the data in different banks could not be linked with one another.

Another participant would have added an obligation for all service providers to report to the user the data they had collected on them. He would have also included all the health companies to service providers that a user could check through a data bank.

One participant considered all the options confusing, including the prevailing situation with present data tracking practices, saying that she “would regard them as absurd if she hadn’t got used to them”. She doubted that the second option could be technically feasible. She also criticised all the options for being commercially orientated and reminded us that originally the internet was not a commercial space. This comment also reminded us researchers how adjusting to the present neoliberal online environment can narrow the ability to imagine other kinds of systems. Furthermore, the idea of a ‘non-commercial internet’ is not only utopian; the BBC, for example, as well as other public service actors, have already taken initiatives to build a “public service internet” (Building a Public Service Internet, BBC Research & Development; Nikunen & Hokka, 2020, see also Fuchs, 2018).

In general, there was a strong tendency among our respondents, corresponding with findings by Selwyn and Pangrazio (2018), that the burden of protecting online privacy should not be left to the user alone. When choosing the second option, many participants explained that that they would need some reliable party to take care of privacy protections on their behalf. Some of the youngest and oldest participants considered it particularly unfair that they were left to personally take care of protecting their personal data against parties they had not even realised were tracking their actions. This is noteworthy, as the traditional understanding of digital literacy places pressure on the individual and underlines the skills that the user should learn to protect herself.

Our workshops also showed that when people with average ICT-knowledge were shown in practice how data-gathering practices work and taught how and why their data is sold by third parties, they were perfectly capable of forming an opinion, and few of them even developed new ideas based on the three options about how they would like data gathering to be organised and regulated. Similar to the results of a study by Kennedy et al. (2017), many respondents thought that they would need more information on how this system works and that there should be more

public discussion on data collection. Offering practical knowledge on data-gathering practices is a good starting point for increased data infrastructure literacy that, in turn, could help average users/citizens better participate in the discussion about the conditions of datafication. As Gray et al. (2018: 9) suggest:

Drawing attention to the politics and making of data and data infrastructures could open up new sites of contestation and controversy as well as creating opportunities for new forms of mobilization, intervention and activism around what they account for. [...] Gaining a sense of diversity of actors involved in the production of digital data (and their interests, which may not align with the providers of infrastructures that they use) is crucial when assessing not only the representational capacities of digital data but also its performative character and role in shaping collective life.

The results mentioned above reassert the ideas of a ‘welfare data society’. Users feel insecure and burdened by the requirement that is built into online environments requiring all users to be solely responsible for her/his own safety against the large platforms or the third parties whose actions were mostly hidden from a user. They long for help from some kind of organisation or institution that they could trust. Fairer user environments can be achieved along two paths that a ‘welfare data society’ should offer: (1) practical digital literacy education to help people grasp the ways in which data gathering works and (2) more analytical data infrastructure education that would help people understand, discuss, and even demand new options to the present situation in which global giants monopolise the online user environment.

In sum, there is a clear need for better data infrastructure literacy so that average users and citizens can be capable of having a political discussion on the ethical aspects of datafication and appropriate regulation. Raising awareness through workshops is effective for small groups, but they are very time-consuming. At the same time, it is urgent to raise general awareness of datafication practices, as a growing number applications that use personal data are continually developed. As Selwyn and Pangrazio (2018) have proposed, there is a strong need for more structural, large-scale solutions to raise the level of digital and data infrastructure literacy. In Finland, one of the major actors in digital literacy education is the Finnish Broadcasting Company YLE. In the next section I analyse the YLE Learning’s content and the actions they have taken in pursuit of

raising data infrastructure literacy and discuss whether European public service media could play a major role in achieving this goal.

YLE LEARNING AS A CONTENT PROVIDER FOR DIGITAL AND DATA INFRASTRUCTURE LITERACY

The YLE Learning (Oppiminen)¹ editorial staff includes an executive editor, a producer, a subeditor, a community manager, and four journalists. As part of the general organisational structure of YLE, it is part of YLE's Creative Content Unit. For this chapter I have interviewed YLE Learning's producer Anna-Leena Lappalainen, with whom we cooperated during the project. According to Lappalainen,

YLE Learning's main task is to promote lifelong learning. It covers categories ranging from digital and media skills, learning skills, school environment, well-being and human relationships, to how society and economics works, and how to develop oneself as a citizen. We approach our topics in an experimental and exploratory spirit.

Unlike many of its European public service counterparts such as BBC Learning² or NRK Skole,³ YLE Learning is not focused on providing content for schoolchildren but to citizens of every age in the spirit of lifelong learning. YLE Learning provides feature articles, educational videos, and quizzes. It has provided digital skills education for several years already and has been producing practical 'digital skills training' content since 2016. In this way, YLE fulfils the traditional public service mission: it provides universal access to education on digital literacy and in this way seeks to empower all kinds of citizens so that they might better cope in a digitalised environment, even if their background education had been left wanting in this respect.

During our cooperation, YLE Learning took datafication as one of their major topics. The decision was grounded in our joint preliminary workshops and YLE's own user workshops in which participants expressed a strong interest in datafication as a journalistic topic. From August 2019

¹<https://yle.fi/aihe/oppiminen>.

²<https://www.bbcstudioslearning.com/index.html>, <https://www.bbc.co.uk/programmes/articles/37gYmkZ17J23P5cxFSL7Q9W/about-the-learning-zone>.

³<https://www.nrk.no/skole/>.

to May 2020, YLE Learning produced seven exploratory pieces that shed light on datafication from different perspectives.

When looking at the seven pieces by YLE Learning analytically, they can be divided into those that support digital literacy and those that could increase data infrastructure literacy. The first group mainly comprises quizzes or short informational packages. In terms of digital literacy (Iordache et al., 2017: 23), they teach users the operational, technical, and formal skills related to digital use and provide guidance in the analysis and evaluation of digital content—a skill that is considered central to digital literacy but also a necessary step in gaining data infrastructure literacy (Büchi et al., 2017; Gray et al., 2018). The second group comprises generally lengthy articles that provide detailed analysis of their topics. Those articles are relevant content for increasing data infrastructure literacy: they help in providing understanding of the present political, social, and economic situation, the “actors involved in the production of data” and their interests, and how digital data now shape everyday life (Gray et al., 2018).

Most of YLE Learning’s datafication pieces are published as pairs, so one piece gives practical advice while the other offers deeper insight into the matter. For example, the first and, so far, the most popular piece is a feature article of an ordinary young woman who tests the GDPR for her own online data. The article explains how she requested that fifteen companies and organisations, from Airbnb to the city of Lahti, provide her with the personal data that they have on her, and explains, in a thriller-like narrative, how each company responded. The article also includes a fact box on how and why companies and organisations gather personal data and what kind of rights the GDPR grants to a private citizen regarding their data. The first story is linked with a second, more practical piece that explains how one can make a request for his or her personal data based on GDPR regulations.

Another pair of YLE Learning’s datafication pieces is a quiz and an educational article related to digital footprints. The quiz, named “What kind of a trace do you leave online?” is made up of questions like, “Do you switch off location data from your mobile phone if you are not using an application that needs it?” or “Have you changed your privacy settings to correspond with your needs in the applications you use?” The quiz has a somewhat similar approach to the user workshops in our research in that it asks the user about her privacy settings. After each answer, YLE’s digital footprint quiz provides a short explanation of why this is important and what option would be useful in light of privacy protections. The quiz is

linked with an educational article that provides nine grounded tips on how to minimise the amount of personal data one shares online.

When we believe that the ability to manage digital footprints is an essential part of digital equality (see Micheli et al., 2018), this kind of accessible content may be quite valuable when trying to raise broad awareness of how to protect personal data. In particular, quizzes may act as effective routes taken to raising awareness of data-gathering practices—such as checking one’s own privacy settings and data, as in our user workshops—though, in face-to-face workshops, users are provided with the opportunity to ask more questions. In addition, the practical tips offered in educational pieces will certainly provide a few more digital skills that will help individuals protect their privacy online.

The fifth article has a slightly different approach as it sheds light on YLE’s own data-gathering practices. The article begins by describing what technically happens when the user opens this webpage and how cookies start to gather data about his or her movements on the site. The article explains which data YLE gathers, why it gathers data, and for what purposes. It also explains that if the user reads the article through Facebook, then Facebook will obtain data about that visit to YLE’s site and use it according to Facebook’s own privacy rules. Again, it expounds the idea that if a YLE news story includes a tweet, Twitter also obtains some user data. The article reveals, on YLE’s behalf, how and why media companies gather user data. This helps the user to understand the now prevalent practices of datafied media with the aim of improving their data infrastructure literacy.

The sixth and seventh articles are again connected to one another. First, YLE Learning has published a nine-minute-long video on the topic “The Internet wants to know everything about you—why should you bother to take interest?” The video features Laura Kankaala, an information security expert who is also known from YLE’s TV series *Team Whack*, in which three ‘white hat hackers’ demonstrate through different case studies how easy it is to hack someone’s personal data. In the video, which also illustrates its points using actors and storytelling, Kankaala explains in detail the many aspects of datafication: why personal data is valuable, how algorithms work through profiling, how algorithms try to make users addicted to social media content, how they may even expose users to political propaganda, and so on. In the end, she urges everyone to take a critical stance towards online incitements and take care of their personal privacy. The seventh article is a profile of Laura Kankaala, in which she also describes

what everyone should take into account when using social media and other online applications. The video sheds light on data-gathering practices but also on the underlying models and the thinking behind data gathering, offering insight into the “politics and making of data” (Gray et al., 2018: 9) and underlying data infrastructures. Content-wise, YLE Learning has much to offer in its attempt to raise the general level of data infrastructure literacy.

Like the first story of the young woman tracing her personal data, the articles on datafication explained by Laura Kankaala have been widely shared on Facebook, and both have managed to reach many readers. In addition, the article providing information on how to make a request for one’s own data based on the GDPR was also very widespread. YLE’s article on data-gathering practices was not so popular among average users but, according to producer Anna-Leena Lappalainen, has gained a lot of positive attention on LinkedIn among IT professionals. Furthermore, Lappalainen noted that unlike regular news articles, YLE Learning’s articles have a long lifespan and are typically found through search engines long after publication by people looking for information on digital skills, digital media, and datafication.

Lappalainen has admitted that datafication is a complicated issue and it is not easy to find story angles that would make datafication interesting to the average reader. However, along with the central PSM ideas of egalitarianism and universalism (Hokka, 2018; Brevini, 2013), they also try to reach those people who are not interested in datafication in the first place and/or have limited education on the subject that could help them understand what datafication and data gathering mean in practice. Yet, what YLE Learning has noticed through their work is that certain storytelling techniques help make datafication a more comprehensible topic. Datafication needs to be linked to everyday life in a very concrete way and the article has to offer something that seems useful, not just something interesting. It helps to explain datafication through an individual and personal perspective, such as in the story of the young woman who requested her own data or through the perspective of some fairly well-known character such as ‘white hat hacker’ Laura Kankaala. Naturally, quizzes that reveal something about the user also pique readers’ interest. However, Lappalainen noted that even though many people, such as the respondents in the user workshops, claim that mere information is enough to get their attention, reader statistics from YLE Learning show that in real terms, sharing pure facts does not induce average readers to get to know

more about datafication; most need some kind of journalistic “kicker” to get started. The insights of YLE Learning’s journalists correspond with previous journalism audience research (Costera Meijer, 2012) in which news readers not only valued journalism that improved the ‘quality’ of their lives but also innovative narrative forms that would increase the pleasure of reading.

Still, when looking at reader statistics at the level of the whole population, the effect of YLE Learning is probably still fairly modest. YLE’s online site,⁴ where YLE Learning’s content is published, reaches 37 per cent of Finns over fifteen-years-old weekly (Nokela et al., 2019), which is impressive when compared to many other European public service media (Schulz et al., 2019: 13). But the average number of readers of YLE Learning’s content is lower, approximately 4 per cent of the Finnish population. The user workshops from our research project indicate that datafication as a topic interests mainly those people who are already somehow familiar with the subject. Despite the positive outcomes of YLE Learning, the question is, what measures could be taken to raise data infrastructure literacy to the level that would make possible a well-informed political discussion on datafication and the democratic decision-making that arises from that newly gained knowledge?

The answer partly lies in what YLE Learning already does. YLE is well-connected with adult education institutes, community high schools and libraries, and they also advertise their content to grammar and high school teachers. In this way, the content helps build digital literacy, and data infrastructure literacy gradually spreads beyond the regular readers of YLE’s website. Furthermore, unlike commercial media, YLE, as a public service media organisation, is free to talk about present data-gathering practices, as their financial stability is not dependent on them unlike more profit-oriented media companies. However, the spread of educational and journalistic content related to datafication currently reaches only those who are seeking out such information. The question remains of how to reach the majority who have trouble allocating the time and/or effort to understanding datafication as a phenomenon with the tremendous effects it has on the development of societies. While public service media are clearly able to produce material that could increase data infrastructure literacy and be an important part of the solution, there is still a need for more coordinated cooperation between different kinds of public organisations

⁴<https://yle.fi/>.

and educational institutions. We clearly need more structural, large-scale solutions for increasing the level of digital literacy (Selwyn & Pangrazio, 2018) and data infrastructure literacy, but the experiences from both user workshops and YLE Learning show that no single actor will manage that mission on their own.

CONCLUSION

Work in the field of critical data studies has recently made great strides in highlighting the many downsides of datafication: surveillance capitalism and dataveillance (Zuboff, 2015; Andrejevic, 2019; Lee, 2019), data colonialism (Couldry & Mejias, 2018; Ricaurte, 2019), data mining, digital footprints and digital traces (Kennedy et al., 2017; Breiter & Hepp, 2018; Micheli et al., 2018), anxieties caused by datafication (Ruckenstein & Granroth, 2019; Lupton, 2019) or the power that algorithms and automatic decision-making possess (Andrejevic, 2020). This remains, however, a work in progress as developing technologies and new products and systems will always force us to confront new ethical dilemmas. Yet, critical data studies should also seek solutions to those dilemmas. That work has already started as many researchers develop new methods and practices that aim to help users as citizens protect their privacy and to look for ways to handle data so that it would benefit users more than it does now (Selwyn & Pangrazio, 2018; Jarke, 2019; Pybus et al., 2015; Kennedy & Moss, 2015; Markham, 2020). In my article, I also attempt to take part in finding solutions for the problems of datafication. As datafication inevitability proceeds, it should be asked, what kind of data society takes care of all citizens' wellbeing and treats citizens in a fair way. How could a datafied society become a welfare data society?

In a welfare data society, the rights and wellbeing of citizens are strengthened through education: by increasing the level of digital and data infrastructure literacy the user/citizen knows her rights and is capable of using them. European countries have already made efforts to improve digital literacy in schooling, but education on digital literacy and especially data infrastructure literacy should also reach those who have not gained this kind of education or whose knowledge is outdated. In my chapter I propose that public service media should also reinforce citizenship through education in an age of datafication.

The EU has taken a leading role in regulating data gathering and granting European citizens the opportunity to manage and control the use of

their personal data. However, for the regulation to be as effective as it should, European citizens need to be more aware of the practices and outcomes of data collection. The results from the user workshops in this study showed how even highly educated users often do not know the amount and type of personal online data about them that is gathered by many kinds of private companies and institutions. Instead, half of the participants were astonished, confused, shocked, and even angry when they realised how much personal data different online services have on them and how detailed it is. Taking care of online privacy requires a certain level of digital literacy that most people do not possess. This leads to digital inequality at the level of the individual when only those users with a fairly good technological education are able to control their digital footprints and protect their privacy online. But, more importantly, it results in an imbalanced and unfair power structure between the ‘Big Tech’ companies and users.

To strengthen the position of users and citizens and to transform the data infrastructures so that they may be fairer towards users, citizens should be informed about the present legal conditions of data gathering so that they might gain reasonable understanding of the political, societal, and cultural consequences of datafication. While this may sound ambitious, our workshops demonstrated that on average, people are capable of forming a thoroughly considered opinion on fair data-gathering practices. Furthermore, they were able to discuss and develop “alternative data regimes and practices” (Kennedy & Moss, 2015) after being introduced to data collection in practice.

In Finland, the public service broadcaster YLE has taken on an educational role as far as the subject of datafication is concerned. The results from our cooperation with YLE Learning show that public service media already possess inventive means through which different kinds of users can be reached—even those users who are not learning digital skills at school or university. Still, more controlled cooperation is needed among different public institutions to increase the number of people who can access the necessary information.

At the same time, work and solutions to increase the level of data infrastructure literacy cannot be left to the cooperation of national institutions only. The average user faces an online environment that is fundamentally global. If we really want to support citizens’ right to be informed and, therefore, able to value the conditions of datafication, there needs to be European-wide cooperation. European public service media organisations

could work together and bring to the fore topical issues related to datafication so that the social and political questions of datafication would not only be discussed by political and academic elites, but by the public at large as well. Possibly, through European-level discussion, new data regimes and practices, such as the data banks that participants preferred, this proposal could be taken forward.

It should also be discussed whether the EBU (European Broadcasting Union) should actively encourage the European PSM organisations to integrate education in digital literacy and data infrastructure into their regular content. At the very least, the EBU should actively support fair and transparent data collection and data use by European PSM companies—and not just to advise PSM organisations to benefit from their user data, as their AI and Data Initiative seems to do.⁵ Only through raising awareness of current data mining practices and their threat to privacy and democracy will citizens be capable of imagining and insisting on new and possible options for the present online environment that GAFAM corporations dominate. The EBU could also take an active role in supporting the initiatives that some PSM organisations have already begun to put into practice, such as the BBC's public service internet model.

If the EU wants its citizens to appreciate and effectively use the GDPR and have more control over their rights to their personal data—or even judge companies and institutions by their ethical standards in data gathering—it should take an active role in increasing the level of data infrastructure literacy among citizens. As our experiences from the user workshops demonstrate, education cannot be left to schools and universities alone, as all kinds of users, and users of every age, need help in gaining the necessary digital skills and data infrastructure literacy. Public service media must also be considered as part of the solution.

REFERENCES

- Andrejevic, M. (2011). Public service media utilities: Rethinking search engines and social networking as public goods. *Media International Australia*, 146(1), 123–132. <https://doi.org/10.1177/1329878x1314600116>
- Andrejevic, M. (2019). Automating surveillance. *Surveillance and Society*, 17(1–2), 7–13. <https://doi.org/10.24908/ss.v17i1/2.12930>
- Andrejevic, M. (2020). *Automated media* (1st ed.). Routledge.

⁵ <https://www.ebu.ch/aidi>.

- Breiter, A., & Hepp, A. (2018). The complexity of datafication: Putting digital traces in context. In A. Hepp et al. (Eds.), *Communicative figurations, transforming communications: Studies in cross-media research* (pp. 387–405). Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-65584-0>
- Brevini, B. (2013). *Public service broadcasting online. A comparative European policy study of PSB 2.0*. Palgrave Macmillan.
- Büchi, M., Just, N., & Latzer, M. (2017). Caring is not enough: The importance of Internet skills for online privacy protection. *Information, Communication & Society*, 20(8), 1261–1278.
- Carr, W., & Kemmis, S. (2009). Educational action research: A critical approach. In S. E. Noffke & B. Somekh (Eds.), *The SAGE handbook of educational action research* (pp. 74–84). Sage.
- Constantiou, I. D., & Kallinikos, J. (2015). New Games New Rules: Big Data and the Changing Context of Strategy. *Journal of Information Technology*, 30(1), 44–57. <https://doi.org/10.1057/jit.2014.17>
- Costera Meijer, I. (2012). Valuable journalism: A search for quality from the vantage point of the user. *Journalism*, 14(6), 754–770.
- Couldry, N., & Mejias, U. A. (2018). Data colonialism: Rethinking Big Data's relation to the contemporary subject. *Television & New Media*, 20, 1–14. <https://doi.org/10.1177/1527476418796632>
- Dencik, L., Hintz, A., & Cable, J. (2016). Towards data justice? The ambiguity of anti-surveillance resistance in political activism. *Big Data & Society*, 3, 1–12. <https://doi.org/10.1177/2053951716679678>
- DR's public-service redecorelse. (2019). <https://www.dr.dk/om-dr/fakta-om-dr/dr-i-2019/dr-i-2019-ny-strategi-massiv-valgdaekning-og-et-meget-mere-digitalt-dr-0>
- Enerhaug, R. F. (2019). Allmennkringkasterregnskapet. Statistikk 2019. <https://www.nrk.no/informasjon/her-er-nrks-rapporter-for-2019-1.14977138>
- European Commission/EACEA/Eurydice. (2019). Digital education at school in Europe. *Eurydice report*. Publications Office of the European Union. https://eacea.ec.europa.eu/national-policies/eurydice/content/digital-education-school-europe_en
- Fuchs, C. (2018). The online advertising tax as foundation of public service internet. *A CAMRI extended policy report*. University Westminster Press. <https://doi.org/10.16997/book23>
- Gangadharan, S. P. (2017). The downside of digital inclusion: Expectations and experiences of privacy and surveillance among marginal Internet users. *New Media & Society*, 19(4), 597–615.
- Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, 5, 1–13. <https://doi.org/10.1177/2053951718786316>
- Gurumythy, A., & Bharthur, D. (2018). Democracy and the algorithmic turn. *Sur International Journal on Human Rights*, 15(27), 39. <https://libproxy>.

- tuni.fi/login?url=https://search-proquest-com.libproxy.tuni.fi/docview/2161599967?accountid=14242
- Ha, V., Al Shaar, F., Inkpen K., & Hdeib, I. (2006). An examination of user perception and misconception of internet cookies. *CHI 2006 extended abstracts on human factors in computing systems*. Association for Computer Machinery, 833–838. <https://doi.org/10.1145/1125451.1125615>
- Hänninen, S., Lehtelä, K.-M., & Saikkonen, P. (2019). Introduction: The Nordic welfare state as a state of civilization. In *The relational Nordic welfare state* (pp. 1–13). Edward Elgar Publishing.
- Helberger, N. (2016). Policy implications from algorithmic profiling and the changing relationship between newsreaders and the media. *Javnost—The Public*, 23(2), 188–203.
- Hokka, J. (2018). Towards nuanced universality: Developing a concept bible for public service online news production. *European Journal of Communication*, 34(1), 74–87. <https://doi.org/10.1177/0267323118810862>
- Holmwood, J. (2000). Three pillars of welfare state theory. T.H. Marshall, Karl Polanyi and Alva Myrdal in defence of the national welfare state. *European Journal of Social Theory*, 3(1), 23–50.
- Iordache, C., Mariën, I., & Baelden, D. (2017). Developing digital skills and competences: A quick-scan. Analysis of 13 digital literacy models. *Italian Journal of Sociology of Education*, 9(1), 6–30.
- Jarke, J. (2019). Open government for all? Co-creating digital public services for older adults through data walks. *Online Information Review*, 43(6), 1003–1020.
- Jensen, C., Potts, C., & Jensen, C. (2005). Privacy practices of internet users: Self-reports versus observed behavior. *International Journal of Human-Computer Studies*, 63(1–2), 203–227. <https://doi.org/10.1016/j.ijhcs.2005.04.019>
- Kangas, O., & Kvist, J. (2013). Nordic welfare states. In *Routledge Handbook of the Welfare State* (pp. 148–161). Routledge.
- Kennedy, H., Elgesem, D., & Miguel, C. (2017). On fairness: User perspectives on social media data mining. *Convergence*, 23(3), 270–288.
- Kennedy, H., & Moss, G. (2015). Known or knowing publics? Social media data mining and the question of public agency. *Big Data & Society*, 2, 1–11. <https://doi.org/10.1177/2053951715611145>
- Lee, C. S. (2019). Datafication, dataveillance, and the social credit system as China's new normal. *Online Information Review*, 43(6), 952–970.
- Lehtiniemi, T. (2017). Personal data spaces: An intervention in surveillance capitalism? *Surveillance & Society*, 15(5), 626–639.
- Lehtiniemi, T., & Ruckenstein, M. (2019). The social imaginaries of data activism. *Big Data & Society*, 6, 1–12. <https://doi.org/10.1177/2053951718821146>
- Lindman, J., & Kuk, G. (2015). From open access to open data markets: Increasing the subtractability of open data. *48th Hawaii international conference on system sciences*. VIELÄ OSOITE

- Lupton, D. (2019). 'It's made me a lot more aware': A new materialist analysis of health self-tracking. *Media International Australia*, 171, 66–79. <https://doi.org/10.1177/1329878X19844042>
- Luzak, J. A. (2014). Privacy notice for Dummie? Towards European guidelines on how to give "clear and comprehensive information" on the cookies' use in order to protect the internet users' right to online privacy. *Journal of Consumer Policy*, 37, 547–559.
- Mai, J.-E. (2006). Big data privacy: The datafication of personal information. *The Information Society*, 32(3), 192–199.
- Markham, A. N. (2020). Taking data literacy to the streets: Critical pedagogy in the public sphere. *Qualitative Inquiry*, 26(2), 227–237.
- Mayer-Schonberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. John Murray.
- Micheli, M., Lutz, C., & Büchi, M. (2018). Digital footprints: An emerging dimension of digital inequality. *Journal of Information, Communication and Ethics in Society*, 16(3), 242–251.
- Nikunen, K., & Hokka, J. (2020). Welfare state values and public service media in the era of datafication. *Global Perspectives*, 1(1), 12906. <https://doi.org/10.1525/gp.2020.12906>
- Nokela, S., Rouvainen, O., & Hyvärilä, A. (2019). YLE's customer relationship. YLE's smart data and audience insight. https://drive.google.com/file/d/1gGZUpAU5PzPNrJ9CR-pf9WA4nqku_VQP/view
- Park, Y. J. (2013). Digital literacy and privacy behavior online. *Communication Research*, 40(2), 215–236.
- Park, Y. J., Chung, J. E., & Shin, D. H. (2018). The Structuration of Digital Ecosystem Privacy and Big Data Intelligence. *American Behavioral Scientist*, 62(10), 1319–1337. <https://doi.org/10.1177/0002764218787863>
- Petersen, C. (2019). User-focused data sharing agreements: A foundation for the genomic future. *JAMIA Open*, 2(4), 402–406. <https://doi.org/10.1093/jamiaopen/ooz043>
- Pybus, J., Coté, M., & Blanke, T. (2015). Hacking the social life of Big Data. *Big Data & Society*, 2, 1–10. <https://doi.org/10.1177/2053951715616649>
- Ricaurte, P. (2019). Data epistemologies, the coloniality of power, and resistance. *Television & New Media*, 20(4), 350–365.
- Ruckenstein, M., & Granroth, J. (2019). Algorithms, advertising and the intimacy of surveillance. *Journal of Cultural Economy*, 13(1), 1–13. <https://doi.org/10.1080/17530350.2019.1574866>
- Ruohonen, J., & Leppänen, V. (2017, 26 December). Whose hands are in the finish cookie jar? *Proceedings—2017 European intelligence and security informatics conference*, 2017, 127–130. <https://doi.org/10.1109/EISIC.2017.25>
- Schreurs, K., Quan-Haase A., & Martin, K. (2017). Problematising the Digital Literacy Paradox in the Context of Older Adults' ICT Use: Aging Media

- Discourse and Self-Determination. *Canadian Journal of Communication*, 42(2). <https://doi.org/10.22230/cjc.2017v42n2a3130>
- Schulz, A., Levy, D. A. L., & Nielsen, R. K. (2019). Old, Educated, and Politically Diverse: The Audience of Public Service News. Reuters Institute. https://www.mediaguru.cz/media/9835/the_audience_of_public_service_news_final.pdf
- Selwyn, N., & Pangrazio, L. (2018). Doing data differently? Developing personal data tactics and strategies amongst young mobile media users. *Big Data & Society*, 5, 1–12. <https://doi.org/10.1177/2053951718765021>
- Susha, I., Rukanova, B., Gil-Garcia, J. R., Tan, Y.-H., & Hernandez, M. G. (2019). Identifying mechanisms for achieving voluntary data sharing in crosssector partnerships for public good. *Proceedings of dg.o 2019, 20th annual international conference on Digital Government Research*. <https://doi.org/10.1145/3325112.3325265>
- Syvrtsen, T., Enli, G., Mjøs, O. J., & Moe, H. (2014). The media welfare state. In *Nordic media in the digital era*. University of Michigan Press.
- Taylor, L. (2016). The ethics of big data as a public good: Which public? Whose good? *Philosophical Transactions of the Royal Society A*, 374, 20160126. <https://doi.org/10.1098/rsta.2016.0126>
- Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, 4, 1–14. <https://doi.org/10.1177/2053951717736335>
- THL Open Data. (2022). <https://thl.fi/en/web/thlfi-en/statistics/statistical-databases/open-data>
- Turow, J. (2011). *Daily you. How the new advertising industry is defining your identity and your worth*. Yale University Press.
- van Dijk, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197–208.
- Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics*, 12(3), 9–26.
- Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30, 75–89.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

