Article:
Steedman, R., Kennedy, H. and Jones, R. (2020) Complex ecologies of trust in data practices and data-driven systems. Information, Communication & Society, 23 (6). pp. 817-832. ISSN 1369-118X

https://doi.org/10.1080/1369118x.2020.1748090

Routledge
Taylor & Francis Group

Information, Communication & Society

ISSN: 1369-118X (Print) 1468-4462 (Online) Journal homepage: https://www.tandfonline.com/loi/rics20

Complex ecologies of trust in data practices and data-driven systems

Robin Steedman, Helen Kennedy & Rhianne Jones

To cite this article: Robin Steedman, Helen Kennedy & Rhianne Jones (2020) Complex ecologies of trust in data practices and data-driven systems, Information, Communication & Society, 23:6, 817-832, DOI: 10.1080/1369118X.2020.1748090

To link to this article: https://doi.org/10.1080/1369118X.2020.1748090

Published online: 08 Apr 2020.

INFORMATION, COMMUNICATION & SOCIETY
2020, VOL. 23, NO. 6, 817–832
https://doi.org/10.1080/1369118X.2020.1748090

Routledge
Taylor & Francis Group

Complex ecologies of trust in data practices and data-driven systems

Robin Steedman*[a], Helen Kennedy[a] and Rhianne Jones[b]

[a]Department of Sociological Studies, University of Sheffield, Sheffield, UK; [b]BBC R&D, Salford, UK

ABSTRACT

Trust in data practices and data-driven systems is widely seen as both important and elusive. A data trust deficit has been identified, to which proposed solutions are often localised or individualised, focusing either on what institutions can do to increase user trust in their data practices or on data management models that empower the individual user. Scholarship on trust often focuses on typologies of trust. This paper shifts the emphasis to those doing the trusting, by presenting findings from empirical research which explored user perspectives on the data practices of the BBC. These findings challenge the assumption that localised or individualised solutions can be effective. They also suggest that conceptualisations of trust in data practices need to account for the complex range of factors which come into play in relation to trust in data and so move beyond the production of typologies. In this paper, we propose the concept of 'complex ecologies of trust' as a way of addressing all of these issues.

ARTICLE HISTORY
Received 22 October 2019
Accepted 6 March 2020

KEYWORDS
Trust; trustworthiness; data; data practices; complex ecologies

Introduction

There is a lot to be concerned about in our digitalised, datafied times: uses and abuses of the data we produce as a result of everyday activities; scandals involving major social media platforms; discriminatory data-driven systems; misinformation and so-called fake news. In this context, trust is both important and elusive. A proliferation of initiatives focusing on trust in data practices and data-driven systems attests to growing interest in this topic. These include policy (e.g., the UK government's National Data Strategy, a main purpose of which is to ensure that data is used in ways that people can trust); conferences (the Association of Internet Researchers' 'Trust in the System' (2019) and 'Digital Trust and Personal Data' (2019) organised by Data For Policy in the UK); and research (the Technology and Trust Initiative at Cambridge University,[1] and Truessec in the EU[2]).

It has been argued that trust is crucial for dealing with uncertain, uncontrollable or risky situations (Sztompka, 1999), a description that could be applied to the data practices (that is, organisations collecting, analysing and sharing data and the outcomes of these processes) that characterise contemporary digital life. How organisations handle personal data is often opaque and beyond the control of most citizens and, in the data ecosystem in

CONTACT Helen Kennedy h.kennedy@sheffield.ac.uk
*Present address: Department of Management, Society and Communication at the Copenhagen Business School

which organisational data practices take place, insecurities are widespread – a recent survey in the UK found high levels of concern about insecurities associated with organisational uses of personal data (Doteveryone, 2018). Other UK-based research has identified what has been described as a 'data trust deficit'. A poll by the Royal Statistical Society (RSS) found that only 6% of citizens report high levels of trust in Internet companies' handling of personal data, and that trust in institutional *data practices* is lower than general levels of trust in those same institutions (2014).

This paper discusses the thoughts and feelings that lie behind the data trust deficit, drawing on empirical, focus group research. We propose the notion of 'complex ecologies of trust' for making sense of what we found and of reported high levels of distrust in data practices and data-driven systems more generally. We argue that the concept of complex ecologies of trust simultaneously contributes to theorising trust and challenges assumptions that trust can be engendered through localised or individualised solutions. In theoretical terms, the concept of complex ecologies of trust moves literature on trust beyond the production of typologies (like generalised and particular trust (Patterson, 1999; Uslaner, 2002)) that characterises much sociological theory, and accounts for the complex range of factors which come into play in relation to trust in data: who is trusting (or not); who or what is being trusted; contexts of trust; and degrees of trustworthiness. By using this concept, we propose that trust is constituted by complex relations across diverse factors, which together engender, maintain or undermine trust in data-driven systems. The complex ecologies of trust that we found in our research also suggest that solutions which are implemented at the level of the particular institution or individual citizen may not be enough; collective or macro-level ways of addressing the data trust deficit may also be needed.

The empirical research on which we base our paper explored perceptions of BBC data practices. We focused on the data practices that accompany the BBC's requirement that its audiences sign into a single account to access core digital services, introduced in 2018, and on possible future data management models. These include gathering demographic and media consumption data in order to personalise recommendations and social media analysis. The rolling out of a sign-in requirement enabled the BBC to deliver elements of personalisation, a key strategic priority in the context of public service media (PSM) across Europe (European Broadcasting Union, 2018), and also to make greater use of audience data. Because trust is an important value for the BBC – 'trust is the foundation of the BBC', its value statements declare (BBC, n.d.-a) – audience views on BBC data practices provide a rich site for exploring questions of trust in data practices and data-driven systems. Focusing on the very specific context of PSM, our paper responds to calls to ground data studies in specific settings in order to 'develop understanding of the material contexts in which datafication has effects' (Kennedy & Bates, 2017, p. 702). At the same time, we believe the complex ecologies of trust that we identified may be found in other contexts, and that this concept can advance understanding of trust in data more broadly. We elaborate on our argument below, after synthesising relevant literature and outlining our methods.

Trust in data practices and data-driven systems

Trust is widely understood to be a socially important, vital and enabling component of everyday life (Sztompka, 1999). Ramírez-i-Ollé cites Confucius, who argued that three things are needed for government: weapons, food and trust (Confucius 2000: 12.7; cited

in Ramírez-i-Ollé, 2019, p. 1). Writing about trust in relationship to democracy, Warren proposes that 'without trust the most basic activities of everyday life would become impossible' (1999, p. 2). For many commentators, trust is especially important in modern, globally uncertain conditions, because trust mitigates the risks that such conditions bring with them. Warren argues that to trust entails accepting 'some amount of risk for potential harm in exchange for the benefits of cooperation' (Warren, 1999, p. 1). Or, as Baier put it, trust is an 'accepted vulnerability to another's possible but not expected ill will (or lack of good will) toward one' (1986, p. 235). Trust involves an orientation toward the future, and it enables us to act, unlike other orientations such as hope, which are less conducive to future action (Sztompka, 1999). Writing about trust as a strategy for dealing with data anxieties, Pink et al. (2018) concur, arguing that trust in data is a feeling that enables people to move on and take action in the future.

In relation to trust in data-driven systems, a number of surveys have addressed aspects of this question in recent years. The Open Data Institute's 2018 survey 'Who do we trust with personal data' focused on which sectors engender the most trust. It found that in the UK, whereas 64% of respondents said they trusted healthcare providers and 57% trusted banking organisations, trust was much lower when it came to other sectors: 10% of respondents said they trusted online retailers and only 2% said they trusted marketing and advertising agencies (Dodds, 2018). As noted above, a poll by the RSS (2014) found that public trust in institutions' data practices is lower than their general levels of trust in the same institutions. Given that only 21% of the UK population tend to trust the media, according to a Eurobarometer report on trust in institutions (2017, p. 15), we might expect trust in media organisations' data practices to be even lower. These polls provide insights into trust in data-driven systems, but some questions remain unanswered, relating to the thoughts and feelings behind low levels of trust, why trust in institutional data practices is lower than trust in institutions more generally, and why data practices are or are not deemed to be trustworthy.

Research into online trust more generally has concluded that trust in the Internet is influenced by personal experience of using it. Drawing on data from an Oxford Internet Institute (OII) Survey, Dutton and Shepherd argue that experience is 'the primary factor shaping trust in the Internet – not prior dispositions shaped by a person's age or gender' (2006, p. 434). Qualitative research by Pink et al. (2018), mentioned above, assessed the individual strategies that people use to navigate uncertainties relating to data in their everyday lives, including at work. The authors found that people build specific, individual routines for managing their data and this helps them to deal with generalised, data-related anxieties, thus also highlighting the importance of experience, in this case in relation to risk and anxiety management.

In literature on trust, distinctions between different types of trust are often made. One is the distinction between particular or strategic trust and generalised or moralistic trust (Patterson, 1999; Uslaner, 2002), in which the former is understood to be directed towards one's own family or group, and the latter towards strangers or collectives like 'fellow citizens' (Li et al., 2018, pp. 4–5). Generalised trust is considered to be conducive to the operation of large-scale and complex networks and institutions (Warren, 1999) like democracies or data ecosystems. Ramírez-i-Ollé (2019) points out that within certain sociological theories of trust, binary distinctions abound, such as low vs high trust, weak vs thick trust, or interpersonal trust vs system/process-based trust. However, the

studies by OII and Pink et al referenced above point to a more dialectic relationship between particular/strategic and generalised trust than such binaries acknowledge. In the OII study, people trust or do not trust in opaque and unfamiliar generalised infrastructures depending on whether they have had particular experiences of them. In the latter case, particular, individual routines are developed in response to anxieties about generalised data-driven systems. These studies hint towards the need for conceptualisations of trust that recognise its complex ecologies.

Ramírez-i-Ollé (2018, 2019) argues that to advance theorisations of trust, we need to move beyond this focus on producing typologies and instead put scepticism at the centre of our thinking. She argues that contrary to the popular belief that trust and scepticism exist in opposition, there is in fact a dialectic relationship between them, much like the relationship between generalised and particular types of trust discussed above. Writing about scientists, she argues that by being sceptical about their own findings – for example by noting methodological flaws or other limitations, rather than claiming authoritative knowledge – scientists enhance the trustworthiness of their field. Thus 'the character of scepticism depends upon the extent and quality of trust' (Shapin, 1994, p. 19; cited in Ramírez-i-Ollé, 2018, p. 2). The same arguments could also be made about data practices.

Data-driven systems are often not trusted for good reason, because they are known either to handle user data in ways which do not inspire trust (as in the Facebook / Cambridge Analytica scandal (Cadwalladr & Graham-Harrison, 2018)) or to discriminate against socially disadvantaged populations (as in reports of the racism embedded in algorithmic criminal justice systems in the US (Angwin et al., 2016)). In such instances, distrust can be appropriate. As Baier puts it, sometimes 'trust-busting can be a morally proper goal' (Baier, 1986, p. 234). People's positions within unequal social structures also inform whether they trust and what they deem to be trustworthy. It has been found that the wealthy and well-educated have higher levels of generalised trust than others (Li et al., 2018; Patterson, 1999). Those whose lives are more insecure trustless, because betrayed trust is more consequential, argues Warren (1999).

Another reason that people may distrust data-driven systems is that these systems are often embedded with the worldviews of their often privileged creators (O'Neil, 2016), as seen in reports by Angwin et al. referenced above, and as widely established by Science and Technology Studies (STS) scholars (e.g., Winner, 1980). Carlson has argued that data-driven systems need to earn the trust of socially disadvantaged populations, for this reason. 'Trust in the system', she argues, 'demands that we begin to infiltrate that system in order to force "it" to incorporate the views and experiences' of those discriminated against as a result of social inequalities (Carlson, n.d.). In other words, systems which do not acknowledge inequalities are unlikely to be deemed trustworthy by people who live with them.

Asking whether data-driven systems are deserving of trust shifts the attention away from those doing the trusting to the trustworthiness of the systems. Trustworthiness has been defined as 'a quality the trustee possesses. It contains an objective dimension which provides ground for the attitude of trust. If a person is trustworthy, then she *can* be trusted i.e., it's rational to trust her' (Stelzer & Veljanova, 2017, p. 11). Trustworthiness has been a focus in literature on public understanding of science, where it has been argued that efforts to increase public understanding as a way of addressing distrust are flawed, in part because of the assumption that greater understanding of science results in greater

trust in it 'remains unproven' (Aitken et al., 2016, p. 713). Aitken et al. argue that to address a trust deficit, the focus should not be on what the public understands, but rather on ensuring that science and related systems are *trustworthy*.

In the context of datafication, a similar argument can be made: the underlying trustworthiness of data-driven systems needs to be improved, not people's understanding of them as data literacy initiatives propose. Yet when solutions for ostensibly more trustworthy systems are suggested, these often focus on what individual institutions or organisations can do. One example can be seen in the services offered by the company Krowdthink, the strapline of which is 'Building privacy, security and trust in digital engagement' (Krowdthink, n.d.). Krowdthink supports its clients to be trustworthy by offering services based on principles such as security- and privacy-by-design, data minimisation, no covert profiling and open business models. These are laudable principles. However, it is important to ask whether such institutional-level solutions are perceived as trustworthy. We address this question in this paper.

Another individualised solution to the perceived data trust deficit that has been proposed is the personal data store (or PDS), which involves individuals personally storing and managing their data. The PDS has received significant attention and financial investment in recent years, and is presented as an ostensibly more trustworthy approach to managing personal data than current models, because it enables individuals to control the processing of, access to and transfer of their personal data (Janssen et al., 2019). Notable examples include Solid[3] led by Tim Berners-Lee, Databox in the UK[4] and services such as digi.me.[5] PDS advocates, such as the international MyData movement, believe that they 'empower individuals by improving their right to self-determination regarding their personal data' (MyData, n.d.). MyData claims that with the PDS, 'the sharing of personal data is based on trust' (MyData, n.d.). However, critics argue that the PDS represents an individualised solution. Lehtiniemi and Ruckenstein state that 'the MyData vision relies on the ethical principle of "human self-determination", treating the individual as an autonomous subject with inalienable rights and liberties' (2019, p. 6). Such polarised views of the PDS suggest that there is a need to explore empirically how solutions targeting individual citizens are perceived by the people that they target, and whether they are seen to adequately address the complex issues relating to data trust that have been identified. We attempt to do this in this paper.

This paper builds on the literature discussed above. It takes account of who is trusting (or not); who or what is being trusted, deserving or undeserving of trust; types and degrees of trust; trust/scepticism relationships; and degrees of trustworthiness. It explores these issues empirically, examining public perceptions of data practices, including some of the proposed solutions to the data trust deficit. It also considers whether social inequalities inform perceptions of the trustworthiness or otherwise of data practices. Below, we outline our methods and discuss our findings.

Research design: context and methods

In the context of public service media (or PSM), trust is an important value, as already noted. PSM have been described as 'islands of trust' in an otherwise self-serving media landscape (Bardoel & d'Haenens, 2008). However, PSM data practices potentially challenge trust, in part because datafication is driven primarily by commercial norms,

which may influence public trust in the data practices of a non-commercial organisation like the BBC (Helberger, 2015; van Es, 2017). As Sørensen and Van den Bulck put it:

> PSM must tread carefully when participating in the user-data-based ecology of third-party servers. They must balance an eagerness to remain relevant with government requirements and competitors' anxieties, while nourishing audience relationships built on trust and on respect for the core values that distinguish PSM from commercial alternatives. (2018, p. 2)

These authors note that there are usually two aspects to trust in PSM: trust that it will be independent of governments and therefore serve democracy, and trust that it will respect audiences and their rights. In relation to the latter, they write, users must trust that personal data 'are not used beyond agreed purposes, and that actions are taken to prevent abuse' (Sørensen & Van den Bulck, 2018, p. 18). This issue of trust emerged as an important theme in our focus group research, which explored perceptions of what happens (and what should happen) to the data that audiences share with the BBC through signing in to access its online services.

We carried out 11 focus groups with 68 participants in two cities in the North of England, the composition of which is summarised in Table 1. In the focus groups, we asked participants about: their media use, especially their use of the BBC's services; their experiences of and attitudes towards signing in to access media services; their feelings about the mining and uses of their personal data by the BBC and other media services and platforms; and their perspectives on possible future data-driven services and data management models. Researching perspectives on data practices is difficult, because they are opaque

Table 1. Participant demographic characteristics.

	Number of participants
Ethnicity	Asian British: 5
	Asian: 2
	Black African: 2
	Black British: 3
	Central and Eastern European: 3
	Mixed/Multiple ethnic groups: 3
	White British: 48
	Other: 2
(Dis)ability	Mild learning disability (such as ADHD, ASD, or Asperger's): 10
	Deaf: 2 (one of which also has a MLD)
	Identified as disabled, but disability not specified: 14
Education level (only highest qualification listed)	No qualifications: 13
	Secondary level qualifications (including GCSE, A Level, apprenticeship, professional qualification, and those currently studying for all but the first of these): 34
	Degree and professional qualification: 12
	Higher degree and professional qualification: 7
	Other/unknown: 2
Employment status	Employed, full time or part time, including self-employed: 18
	Student: 17
	Unemployed: 10
	Retired: 19
	Unknown/other (including carer): 4
Age	16–34: 28
	35–34: 21
	65+: 19
Gender	Women: 41
	Non-binary: 1
	Men: 26

by design and there is widespread confusion about them. We adopted a range of strategies to address this challenge. One of these was to use a cyclical approach. First, we identified what participants already knew about the BBC's data practices. Then, we provided them with information about these practices, adding to or clarifying what they already knew. We drew on publicly available information to do this. Finally, we explored how participants felt about the knowledge they acquired through participation in the focus groups.

Exploring what participants felt about future data-driven services and data management models was a further challenge, as it required them to imagine future scenarios beyond their experience. To address this challenge, we presented participants with concrete examples, including the commonplace data management model through which organisations collect, store, manage and control user data, and the PDS model, discussed above, in which users decide who can access their data, for what purposes and under what circumstances. After discussing these models, we asked participants to share any ideas they had for alternative models, which could include the ability to opt out of data collection entirely. Thus we encouraged participants to imagine future scenarios, but only after discussing concrete examples with them.

Focus groups lasted between one and two hours. They were audio-recorded, transcribed and anonymised, after which audio-recordings were deleted. We gave participants a £20 one4all voucher to thank them for their participation in our research. We secured ethical approval for our research from our university, and complied with GDPR (new EU General Data Protection Regulation) when it came into force half way through data collection. Data were organised and coded in Nvivo, according to pre-determined and emergent themes and categories. In the next section, we discuss what emerged from our analysis.

Complex ecologies of trust in BBC data practices

In our research, we found complex ecologies of trust in BBC data practices. For many participants, feelings about organisational data practices had little to do with the actual organisational data practices about which feelings were expressed. Rather, feelings of trust or distrust in data practices were shaped by participants' views and experiences of the organisation responsible for the data practices. So unlike the RSS survey cited above (2014), which found that public trust in institutions' uses of their data was lower than their general levels of trust in the same institutions, for some of our participants, there was a parallel between the two.

Positive views of the BBC led some participants to trust its data practices. Melissa (25–34, white British, employed, educated to degree and professional qualification level) was proud of the BBC because it is a 'world-renowned service' and therefore she trusted its data practices. She said 'I think the BBC is like probably the least of my worries when it comes to people taking and stealing data.' James (55–64, white British, educated to degree and professional qualification level, employed part-time) also trusted the BBC more than other organisations that gather user data. This was because 'it's an institution that's been around a longer time,' as he put it, and because he personally knew BBC employees. For James, trust in the BBC was both generalised and particular: he saw the BBC as trustworthy because of its longevity as an institution and because of his personal connections with BBC employees.

Likewise, Karen (65–74, white British, disabled, retired) felt that, as a 'well-established' institution, the BBC would be 'above' the problematic data practices of other organisations. She also noted that there had not been any public data scandals involving the BBC, in contrast to other major digital service providers. Sarala (35–44, Asian British, educated to degree level, employed) concurred. She was less worried about the BBC than about social media platforms because of its reputation: 'Maybe because again it's the name, you have trust in the BBC, don't you, because we trust it for everything else like news and giving us the right information and everything.'

In contrast, other participants held less positive views of the BBC, which led them to distrust its data practices. Patricia (35–44, white British, disabled, educated to degree level, unemployed) was concerned about a negative news story about a BBC presenter that had surfaced in the recent past and because of this, she asked 'can you trust the BBC with anything?' Patricia did not trust the BBC to handle her data because of an entirely unrelated incident; her distrust in the BBC as a whole led her to distrust its data practices.

In the focus groups, we showed participants extracts from the BBC's privacy promise at the time of the study (BBC, n.d.-b), which included a promise to protect users' personal data. We asked participants whether they trusted the BBC to uphold this promise. Some did, and others did not. Steven (55–64, white British, educated to secondary level and professional qualification level, retired) did not trust the BBC in general because he thought it was biased and had hidden agendas. He therefore did not trust its privacy statement, declaring 'I don't trust it for one minute. […] everything could be sellable.'

Some participants in our study thought that the BBC was trustworthy, but that it did not have the capacity to uphold its promise. Brian (45–54, British, other ethnic group, educated to higher degree level, retired) said: 'I think I subscribe more to the cock-up theory […] rather than the conspiracy theory. […] I mean are they actually able to look after the data properly?' He thought the BBC privacy statement was made in good faith but that 'there might be [data] leakage by accident.' This comment may explain the RSS survey finding that trust in institutional data practices is lower than their trust in the institutions in general. The institution may be trusted, but anxiety about the safety of the broader data ecosystem may lead people not to trust the institution's ability to engage in responsible data practices.

This was the case for many of our participants. Anxiety about the data ecosystem often meant participants did not trust BBC data practices, regardless of whether they trusted the organisation. High levels of trust in the BBC co-existed with knowledge about data breaches elsewhere, which made participants feel that their online data was not safe. For instance, according to Melissa (mentioned above), 'your data is just never safe online.' Erik (18–24, white, Norwegian, university student) said he trusted the BBC's intent in relation to protecting users' data, but he saw data breaches in other companies as evidence that the BBC could also experience similar problems. This confirms Flyverbom's (2017, p. 73) assertion that data breaches 'amplify the erosion of trust in internet companies and digital infrastructures.' Nicole (45–54, white British, educated to secondary level, employed) expressed similar feelings to Melissa and Erik. She said:

> banks and just about everybody always tells you that they're going to keep your personal data safe but you see things all the time where it's not safe is it? […] so I don't think they can say that they can keep it 100% safe.

George (75+, white British, educated to apprenticeship level, retired) said he felt the same. He thought that it was important to be 'dead careful' when sharing data online, and this view influenced his online interactions with BBC services.

While a number of participants recalled specific data breaches, these resulted in ill-defined and generalised anxieties about data sharing. As in Pink et al.'s (2018) study, we thus saw a dialectical relationship between the generalised and the particular. Participants felt insecure, but the precise nature of their insecurities was hard to express. In our focus group with an Asian family, two family members, Arjun (35–44, Asian British, educated to degree level, employed) and Sarala (mentioned above) said they were mindful about what they did online and concerned for vulnerable people who may not understand data-driven systems as well as them. When the younger Neha (18–24, Asian British, recent A-level graduate) said that she did not understand how online activity could be dangerous, her older relatives tried to explain, but struggled to articulate the dangers that concerned them. Similarly, Marilyn (75+, white British, disabled, retired) said 'there's always going to be an evil element that will find a way to access your information [online],' but she did not have a clear idea of what this 'evil element' could be.

The entanglement of factors outlined here results in a dialectic and dynamic interplay between trust, scepticism and distrust. Some participants trusted the BBC as a UK-based PBS and trusted its data practices, some distrusted both, and others trusted the BBC but were sceptical about its ability to handle user data securely within a hyper-connected global data ecosystem. Perceptions of the trustworthiness or otherwise of the BBC were both particular and generalised, influenced by personal relationships with BBC employees, longstanding reputation, recent news stories, or a combination of these factors. In the case of James, for example, particular trust – in people who work for the BBC – led to generalised trust in the organisation. Overall, the findings discussed here point to the need to move beyond typologies (like trust/distrust, particular/generalised) and to develop a way of conceptualising trust in data-driven systems that accounts for complexity. We suggest that the concept of complex ecologies of trust is one way of doing this. 'Complex ecologies of trust,' we suggest, takes account of the multiple factors that engender, maintain or undermine trust in data-driven systems, including experience, perception, understanding and feelings as they relate to organisations, services, people and practices.

Building trust in data systems: challenging individual and institutional solutions

Recent surveys and polls, some of which we have discussed above, have concluded that Internet users want more control over what happens to their data, and that giving users more control might lead to a greater degree of trust in data-driven systems. We explored this issue in our study by discussing two different models for managing personal data, as outlined above. We focused on the commonplace current model, which gives organisations control over user data, and a model based on the PDS, because BBC R&D was experimenting with such a model at the time of our research (Thompson & Jones, 2019). We also invited participants to share their own ideas for alternative data management models, but there were no noteworthy responses to this invitation. We therefore focus our discussion here on participants' views of the PDS.

Participants generally agreed that a PDS is better than the current model, because it offers more control to individual users. Participants tended to share Sana's (35–44, Asian British, educated to professional qualification level, employed) view: 'I prefer to be in control of my data and what they're using it for. […] if they're collecting it anyway then I'd rather be the one to tell them […] what they should use it for.' Leyla (16–24, Somali-British, university student) described the PDS model as 'fair' because users have a choice about what data to share without this choice impacting their ability to use a service. Maya (16–24, mixed/multiple ethnic groups, recent university graduate) thought it was 'courteous,' because it requests data (rather than demanding it in exchange for services, which was how she perceived the current model) and explains how the data will be used.

However, while the PDS was perceived as preferable to current data management arrangements, participants were not convinced that it would address all of their anxieties. Concerns about the PDS revolved around two issues: time and security. Participants were worried that managing personal data through the PDS would require significant investments of time by individual users. For Hazel (25–34, white British, Deaf, educated to higher degree level, employed), the transparency of the PDS was 'really attractive,' but being in control of it and managing it 'could be a lot more hassle.' When asked whether he would prefer to manage his PDS himself or entrust it to an intermediary, Andrew (35–44, white British, educated to secondary level, employed) responded, 'it depends on how much time and effort it would take to do it yourself.'

Like Andrew, Estrella (35–44, mixed/multiple ethnic groups, Spanish, disabled, Master's student) was concerned that managing personal data through a PDS could be time-consuming. She felt that 'it's a great idea' but continued 'I would never use it, because that implies time and answering somebody else's questions.' Samantha (18–24, white British, educated to apprenticeship and professional qualification level, employed) worried that users of the PDS model might be 'bombarded with loads and loads of emails' asking for data. Others were also concerned they would not have enough time to devote to managing their PDS well. For Melissa (mentioned above), every time data is requested:

> it's a decision you have to make, it's like weighing it up "oh, what could they do with this? Is this going to be a problem for me?" So you either go through that or you just click on it and say "yeah yeah whatever" which is what we kind of do now, most of us. […] I don't really know what they're doing with my data and I feel like we're not really informed enough, but at the same time I don't have time to read through a load of different like things on T&Cs or whatever.

Melissa worried that as a PDS user, she would not have the time to make thoughtful decisions and this would leave her in a situation similar to current arrangements. When Arjun (mentioned above) said the process of managing a PDS 'has to be completely easy' and 'especially user friendly,' he touched on the same issue. For these participants, the PDS was potentially burdensome in the same way that reading the terms of service of digital platforms has been shown to be (for example by Obar & Oeldorf-Hirsch, 2018).

Most participants felt that making thoughtful decisions about what to do with their data would take time, and that a high number of requests for data from different services would make this difficult. Many acknowledged that they currently do not think carefully when signing up for services because they do not have time to do so. These participants were

critical of contemporary forms of data gathering, but felt that the ubiquity of these practices diminished their ability to act on their own critical thinking and consider their choices carefully. Others, like Michael (18–24, white British, university student) and Maya (mentioned above), felt that the normalisation of data gathering had made them complacent about it. This could be seen as a form of what Draper and Turow (2019) describe as 'digital resignation.' Participants were concerned that the same thing could happen with a PDS. To address this concern, some participants said they would like to see options to deny all requests to access their data in a PDS, to prohibit certain organisations from requesting data, or to set their default response to data requests to 'no' with the ability to change this to yes in certain cases. Thus PDSs were seen to give more control, but individual responsibility for them was not desirable, because it was perceived as time-consuming.

A further concern related to security. Many participants wondered if the PDS was secure, a concern they also expressed about the broader data ecosystem. Because a PDS stores personal data in some form, in the cloud in the case of digi.me, or on physical devices such as Databox, some participants assumed that data would be aggregated in a physical location and were worried about the security of this arrangement. Tereza (18–24, white, Czech, educated to secondary level, employed), for example, was concerned about having all of her personal information on a PDS as opposed to distributed across various services. Tom (65–74, white British, educated to professional qualification level, retired) emphasised that 'you'd have to know it was a secure site' to feel comfortable using a PDS. Others felt that such security was not possible – all 12 members of one focus group did not trust that a PDS could be secure within the current, insecure data ecosystem. Virginia (75+, white British, retired) asked: 'how confident could you be that your data would not be shared?'

Thus the PDS did not allay concerns about unnamed others accessing personal data; most participants were sceptical about this data management model. Our findings suggest that solutions targeting individual users and implemented at the level of the individual organisation do not address concerns about the data ecosystem as a whole. Instead, some participants implied that more macro level solutions may be more effective, and others discussed this explicitly. A small number of participants spoke about the need for a regulatory framework as a solution to the data trust deficit. For example, Brian (mentioned above) saw regulation as a solution to his concerns about PDS data security:

> Supposing they start surreptitiously transferring the data from your databox to their database. I mean what protects you from that? […] The fact that you've got a data box, I mean how does that help you? […] It doesn't help you without, without some legal protection.

For Brian, individual control of personal data does not solve wider data-related problems, because individuals cannot police organisations' misuse of data. Other participants also recognised the value of a strong regulatory framework for data protection. Hazel (mentioned above) trusted the BBC more than global social media platforms because 'they're a UK-based company and they're bound by European legislation.' Also reflecting on BBC data practices, Sana (mentioned above) said, 'there's more control because of GDPR.' These reflections indicate that participants saw a need to go beyond individual and institutional solutions, to imagine instead solutions based on a more

collective notion of 'our data' rather than the individualised notion of 'my data' (Lehtiniemi & Ruckenstein, 2019).

Another way in which participants pointed to the need for collective solutions was by questioning the accessibility of a PDS to all potential users. For instance, Pamela (45–54, white British, no educational qualification, unknown occupation) thought the PDS model 'might be interesting now with kids, the young kids who are at school now; but not for the likes of us.' Above we noted how Arjun and Sarala recognised how inequalities might affect understandings of data-driven systems and PDS experiences. These comments point to a recognition amongst some of our participants that people's positions within unequal social structures may inform their ability to use, and therefore their inclination to trust, a technological solution like the PDS (Patterson, 1999; Warren, 1999).

These comments point to the question of whether social inequalities shape the data trust deficit. In another paper based on this study, we have argued that inequalities relating to age, dis/ability, poverty and their intersections played a role in informing understandings of and feelings about data practices (Authors, in press). Older participants, younger participants with mild learning disabilities and participants with experiences of poverty tended to understand BBC data practices less well than other participants, although many still had strong feelings about them. We did not, however, identify such patterns when it came to questions of trust. Across diverse participant groups, we found some trust, some distrust, some scepticism, and often, dialectic combinations of these things. For this reason, we use the concept *complex* ecologies of trust to characterise our findings – no simple patterns emerged, and so a way of conceptualising trust in data-driven systems that accounts for complexity is needed.

In relation to our participants' reflections about the PDS, we suggest that complex ecologies of trust were at play. Distrust in the data ecosystem led to scepticism about the PDS as a solution to participants' data-related anxieties. Distrust in the data ecosystem meant that efforts to bridge the data trust deficit at the level of the individual organisation would not be enough on their own, for some of our participants. The PDS is a particularised solution, targeted at individual users and implemented by individual organisations, which, alone, does not address generalised anxiety. Participants' awareness of social inequalities, or how socially unequal populations might respond differently to technologies like the PDS, also suggests that individualised solutions may be inadequate. Lehtiniemi and Ruckenstein (2019) suggest that the worldview embedded in the PDS model is individualistic, something our participants implicitly acknowledged in their discussion of their concerns about it, discussions which acknowledged social inequalities. Societal-level solutions like regulation might do more to engender trust, therefore, than these individualised solutions.

Conclusions

We argue that the notion of 'complex ecologies of trust' captures the multiple, interrelated factors that engender, maintain or undermine trust in data practices and data-driven services. In relation to our findings, complex ecologies of trust helps us to make sense of the interplay between trust, scepticism and distrust that we identified in our participants' views of BBC data practices. Within this complex interplay, most participants distrusted the broader data ecosystem in which BBC data practices take place. This meant that efforts to bridge the data trust deficit by an individual organisation or that were targeted at an

individual user were not perceived by our participants to address the causes of their anxieties. Distrust in the data ecosystem led to scepticism about the PDS as a solution to data-related anxieties, in part because distrust was at a general level whereas the PDS solution targets the particular. Furthermore, some participants' awareness of social inequalities and how these might limit the accessibility of a PDS-like model also led them to doubt its efficacy as a solution to the data trust deficit.

The notion of complex ecologies of trust also helps us to make sense of trust as context dependent. In our case study, the context was public service media. Some participants trusted the BBC, in general and in its data practices, because it is a public service broadcaster, and this shaped their perceptions of its trustworthiness. The broader data context caused anxiety for many participants, and this is likely to inform perceptions of data practices in other contexts. Because of this, changes are needed at the macro level of the data ecosystem in order to engender greater trust, not in individual institutions or in relation to individual users. In other words, particularised solutions to generalised problems are unlikely to be effective. We need collective, ecosystem solutions, for example, better regulation of data-driven systems, in order for them to be perceived as more trustworthy. Furthermore, to ensure that socially unequal populations are not disadvantaged by data-driven systems, collective data management models like public data commons or co-operatives (ODI, 2019) may be more effective than personalised alternatives. To understand what members of the public think about these and other possible data management models, further research is needed.

A further complexity in trust in data practices relates to layers of trust – for example, whether participants trust the BBC in general, trust the BBC specifically to manage their data securely, trust the broader data ecosystem, and even whether they trust themselves to manage a PDS carefully and thoughtfully. Combined with the other issues we have discussed in this paper, these layers, point to the need to develop a way of conceptualising trust in data-driven systems that accounts for all of these complexities. The concept of complex ecologies of trust is one way of doing this, as it takes account of the multiple factors discussed here, and how they inform people's experiences, perceptions, understandings and feelings in relation to the trustworthiness or otherwise of data practices and data-driven systems.

Despite widespread distrust, data-driven services carry on flourishing, and what citizens say about what would make them more trustworthy appears to go unheeded. Almost half of respondents to doteveryone's Digital Attitudes survey (2018) agreed that 'it doesn't matter whether they trust organisations with their data online,' because companies do what they want anyway and they, the users, are dependent on the services that these companies offer – Draper and Turow (2019) refer to this as 'digital resignation,' as noted above. The ODI concludes its data trust survey report by stating that an 'overwhelming factor in considering whether or not to share data […] is whether or not people trust the organisation asking for it' (Dodds, 2018). But people do not choose whether or not to share their data, and the suggestion that data is 'shared', with its connotations of giving freely in a context characterised by equal social relations (John, 2013, 2017), is misleading at best. Trust is claimed to be a vital enabling component of social life, and yet in its absence, data-driven social life continues. Misztal draws on her experiences in communist Poland to consider how societies function when trust is absent. She writes that 'the absence of trust does not necessarily entail a lack of social order since order can be sustained by

effective government based on fear' (Misztal, 1996, p. 64, cited in Ramírez-i-Ollé, 2019, p. 2). She goes on to say that we need to ask what kinds of order are supported by trust, but it is more important, perhaps, to consider what kinds of order flourish in its absence.

Disregard for public views about what would make data practices more trustworthy will no doubt perpetuate distrust in the data ecosystem, and this will have consequences for PSM organisations which are trying to work with user data in ways which do not undermine public trust. Established trust in PSM may be challenged by the commercial norms that drive datafication, even if PSM does not monetise data in the same way that commercial organisations do. A further challenge in the complex ecology of trust, then, is to consider whether not mining and storing personal data at all is the only solution for organisations that want their users to trust them.

Funding

This work was supported by the Arts and Humanities Research Council.

Notes

1. https://www.trusttech.cam.ac.uk.
2. https://truessec.eu/.
3. https://solid.inrupt.com/.
4. https://www.databoxproject.uk/.
5. https://digi.me/.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Robin Steedman is a postdoctoral researcher interested in diversity and inequality in the media, based at the Copenhagen Business School.

Helen Kennedy is professor of digital society at the University of Sheffield and director of the Living With Data programme of research (https://livingwithdata.org/).

Rhianne Jones is a research lead at BBC Research & Development. Her research focuses on examining the development and use of data-driven technology and the public interest.

References

Aitken, M., Cunningham-Burley, S., & Pagliari, C. (2016). Moving from trust to trustworthiness: Experiences of public engagement in the Scottish Health Informatics Programme. *Science & Public Policy*, *43*(5), 713–723. https://doi.org/10.1093/scipol/scv075

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). *Machine bias* [Text/html]. Retrieved September 27, 2019, from ProPublica website: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Authors. (in press).

Baier, A. (1986). Trust and Antitrust. *Ethics*, *96*(2), 231–260. https://doi.org/10.1086/292745

Bardoel, J., & d'Haenens, L. (2008). Reinventing public service broadcasting in Europe: Prospects, promises and problems. *Media, Culture & Society*, *30*(3), 337–355. https://doi.org/10.1177/0163443708088791

BBC. (n.d.-a). *BBC values*. Retrieved October 1, 2019, from About the BBC website: bbc.com/about-thebbc//governance/charter

BBC. (n.d.-b). *The BBC privacy promise*. Retrieved October 1, 2019, from Using the BBC website: https://www.bbc.co.uk/usingthebbc/privacy-promise

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Carlson, B. (n.d.). *Indigenous internet users: Learning to trust ourselves*. Retrieved September 27, 2019, from AoIR 2019 Keynote and Plenary Panel website: https://aoir.org/aoir2019/aoir2019program/

Dodds, L. (2018, July 5). *Who do we trust with personal data?* Retrieved February 18, 2019, from Open Data Institute website: https://theodi.org/article/who-do-we-trust-with-personal-data-odi-commissioned-survey-reveals-most-and-least-trusted-sectors-across-europe/

Doteveryone. (2018). *People, power and technology: The 2018 digital attitudes report*. https://doteveryone.org.uk/report/digital-attitudes/

Draper, N. A., & Turow, J. (2019). The corporate cultivation of digital resignation. *New Media & Society*, *21*(8), 1824–1839. https://doi.org/10.1177/1461444819833331

Dutton, W. H., & Shepherd, A. (2006). Trust in the Internet as an experience technology. *Information, Communication & Society*, *9*(4), 433–451. https://doi.org/10.1080/13691180600858606

Eurobarometer. (2017). *Designing Europe's future: Trust in institutions; globalisation; support for the euro, opinions about free trade and solidarity*. European Commission. https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/SPECIAL/surveyKy/2173

European Broadcasting Union. (2018). *Big data initiative: Activity report 2017-18*. https://www.ebu.ch/publications/big-datat-initiative---activity-report-2017-2018

Flyverbom, M. (2017). *Datafication, transparency and trust in the digital domain*. Trust at Risk: Implications for EU Policies and Institutions: Report of the Expert Group "Trust at Risk? Foresight on the Medium-Term Implications for European Research and Innovation Policies (TRUSTFORESIGHT) (pp. 69–84). https://doi.org/10.2777/364327

Helberger, N. (2015). Public service media. Merely facilitating or actively stimulating diverse media choices? Public service media at the crossroad. *International Journal of Communication*, *9*, 1324–1340. https://ijoc.org/index.php/ijoc/article/view/2875/1374

Janssen, H., Cobbe, J., Norval, C., & Singh, J. (2019). Personal data stores and the GDPR's lawful grounds for processing personal data. *Zenodo*, 1–6. https://doi.org/10.5281/zenodo.3234902

John, N. A. (2013). Sharing and Web 2.0: The emergence of a keyword. *New Media & Society*, *15*(2), 167–182. https://doi.org/10.1177/1461444812450684

John, N. A. (2017). *The age of sharing*. Polity.

Kennedy, H., & Bates, J. (2017). Data power in material contexts: Introduction. *Television & New Media*, *18*(8), 701–705. https://doi.org/10.1177/1527476417720034

Krowdthink. (n.d.). Krowdthink. Retrieved October 1, 2019, from https://krowdthink.com/

Lehtiniemi, T., & Ruckenstein, M. (2019). The social imaginaries of data activism. *Big Data & Society*, *6*(1), 1–12. https://doi.org/10.1177/2053951718821146

Li, Y., Smith, N., & Dangerfield, P. (2018). *Social trust: The impact of social networks and inequality* (pp. 1–25). The National Centre for Social Research.

MyData. (n.d.). Homepage. Retrieved September 27, 2019, from MyData.org website: https://mydata.org/

Obar, J. A., & Oeldorf-Hirsch, A. (2018). The biggest lie on the Internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication Society*, *23*(1), 1–20. https://doi.org/10.1080/1369118X.2018.1486870

ODI. (2019, April 15). *Huge appetite for data trusts, according to new ODI researc*h. Retrieved September 30, 2019, from https://theodi.org/article/huge-appetite-for-data-trusts-according-to-new-odi-research/

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy (1st edition)*. Crown.

Patterson, O. (1999). Liberty against the democratic state: On the historical and contemporary sources of American distrust. In M. E. Warren (Ed.), *Democracy and trust* (pp. 151–207). Cambridge University Press. https://doi.org/10.1017/CBO9780511659959.006

Pink, S., Lanzeni, D., & Horst, H. (2018). Data anxieties: Finding trust in everyday digital mess. *Big Data & Society*, *5*(1), 1–14. https://doi.org/10.1177/2053951718756685

Ramírez-i-Ollé, M. (2018). 'Civil skepticism' and the social construction of knowledge: A case in dendroclimatology. *Social Studies of Science*, *48*(6), 821–845. https://doi.org/10.1177/0306312718763119

Ramírez-i-Ollé, M. (2019). Trust, scepticism, and social order: A contribution from the sociology of scientific knowledge. *Sociology Compass*, *13*(2). https://doi.org/10.1111/soc4.12653

Royal Statistical Society. (2014, October 2). *Royal Statistical Society research on trust in data and attitudes toward data use / data sharing*. https://www.statslife.org.uk/images/pdf/rss-data-trust-data-sharing-attitudes-research-note.pdf

Sørensen, J. K., & Van den Bulck, H. (2018). Public service media online, advertising and the third-party user data business: A trade versus trust dilemma? *Convergence*, *26*(2), 421–447. https://doi.org/10.1177/1354856518790203

Stelzer, H., & Veljanova, H. (2017). *Deliverable D4.2 support study: Ethical issues: Periodic activity and management report* (pp. 1–45). https://truessec.eu/sites/default/files/evidence/d4.2_support_study_-_ethical_issues.pdf

Sztompka, P. (1999). *Trust: A sociological theory*. Cambridge University Press.

Thompson, B., & Jones, R. (2019, July 2). *Introducing the BBC Box*. Retrieved September 27, 2019, from BBC R&D website: https://www.bbc.co.uk/rd/blog/2019-06-bbc-box-personal-data-privacy

Uslaner, E. M. (2002). *The Moral Foundations of Trust*. https://doi.org/10.1017/CBO9780511614934

van Es, K. (2017). *An impending crisis of imagination: Data-driven personalization in public service broadcasters* (Media@LSE Working Paper Series).

Warren, M. E. (1999). Introduction. In M. E. Warren (Ed.), *Democracy and trust* (pp. 1–21). Cambridge University Press.https://doi.org/10.1017/CBO9780511659959.001

Winner, L. (1980). Do Artifacts have Politics? *Daedalus*, *109*(1), 121–136.