



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/215892/>

Version: Published Version

Article:

Kennedy, H., Ditchfield, H., Oman, S. et al. (2024) How people connect fairness and equity when they talk about data uses. *Big Data and Society*, 11 (4). ISSN: 2053-9517

<https://doi.org/10.1177/20539517241303162>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

How people connect fairness and equity when they talk about data uses

Helen Kennedy¹ , Hannah Ditchfield¹ , Susan Oman² ,
Jo Bates³ , Itzelle Medina Perea³ , Monika Fraczak¹ ,
and Mark Taylor⁴

Big Data & Society
October–December: 1–15
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517241303162
journals.sagepub.com/home/bds



Abstract

As a mechanism for addressing data-related harms, fairness has been subjected to considerable criticism, seen as failing to acknowledge the power relationships that produce said harms, or as a ‘floating signifier’ devoid of specific meaning. In contrast to fairness, it is argued that equity does a better job of recognising data-related harms. Criticisms such as these emerge in specific cultural contexts and rarely acknowledge everyday understandings of terms and concepts. This paper engages with these criticisms, drawing on research exploring how 112 UK residents perceive data uses in specific public service organisations. We found that participants perceive fairness and equity to be interwoven with each other, a finding which shows that who gets to define what is fair matters and which challenges assumptions about what does and does not constitute thinking and talking data politics. We conclude by proposing that linking fairness with equity can be seen as a kind of everyday data solidarity.

Keywords

Data uses, equity, everyday life, fairness, talking data politics

Introduction

It is well established that ever-increasing uses of data across varied domains of life can result in an array of harms. Identifying mechanisms with which to avert these harms is thus an urgent and ongoing challenge, but opinions differ about proposed responses. For example, the proposition that data-driven systems should be fair, accountable and transparent has been criticised within the field of critical data studies. Fairness, the focus of this paper, is seen as distinct from and dichotomous to equity and justice (e.g. by D’Ignazio and Klein, 2020) which, it is claimed, are more useful tools for recognising and addressing data-related harms (Barabas et al., 2020; Miceli et al., 2022). Furthermore, in the context of data- and AI-driven systems, the adjectives fair and good are seen as ‘infinitely spacious words that any AI system can be squeezed into’ (Kalluri, 2020). In other words, they are what Lévi-Strauss (1987) describes as ‘floating signifiers’ – their meaning is unspecified and this opens them up to multiple, sometimes questionable uses.

These criticisms emerge from specific, often US-based contexts and particular, activist perspectives; whether they hold true in other contexts remains unknown. Furthermore, definitions of the term fairness, and applications of the term in debates about data systems, are primarily derived

from data experts, including developers, practitioners and data activists, rather than members of the public. As such, experiential, everyday perspectives are often overlooked. This is a problem because, as sociologists of everyday life argue, social and political phenomena are shaped by everyday understandings of them (Davies, 2021; Neal and Murji, 2015). Focusing on everyday perceptions of fairness in relation to data can therefore advance the very debates from which public voices are often excluded and make them more inclusive.

Centuries of structural inequity and overlapping systems of oppression mean that some voices are more excluded than others in debates about data (Patel and Jones, 2021). Structural inequity also shapes everyday experiences and perceptions. A further focus of this paper, therefore, is on

¹Department of Sociological Studies, University of Sheffield, Sheffield, UK

²Centre for Machine Intelligence, University of Sheffield, Sheffield, UK

³Information School, University of Sheffield, Sheffield, UK

⁴Sheffield Methods Institute, University of Sheffield, Sheffield, UK

Corresponding author:

Helen Kennedy, Department of Sociological Studies, University of Sheffield, Sheffield, UK.

Email: h.kennedy@sheffield.ac.uk



equity and inequity (hereafter in/equity) and the role these play in people's thinking about fairness and data uses. We understand data in/equity as an element of the broader and more widely discussed umbrella term data justice (Dencik et al., 2019; Taylor, 2017) which is in need of empirical scholarly attention.

We report on qualitative research with 112 UK residents which explored their views on public sector data uses, a phrase we use to refer to data collection, analysis and sharing. We focused on uses of personal data – that is, data 'related to an identified or identifiable person' (EU, 2016) – in welfare, health and public service broadcasting contexts. At the time of our data collection in 2021, public views on data uses in these contexts were under-researched, despite their importance in everyday life. We asked participants what they thought of six specific data uses, whether they thought they were fair or unfair, and what they thought a fair data use looked like. To operationalise our focus on in/equity, we recruited a sample which was inclusive of participants with demographic characteristics on which inequities and associated discrimination are often based, relating to gender, race and ethnicity, age, disability, sexuality, religion, income and education. Eighty-five per cent of the participants were from one and 46% were from two or more minority or disadvantaged groups. We also explored how social inequities intersect (Crenshaw, 1991) in perceptions of data uses.

We argue that in participants' perceptions of data uses, fairness and equity were not as distinct and dichotomous as critical commentators propose. Rather, the terms were interwoven with each other, with fairness linked to concerns about in/equity. We therefore suggest that fairness is not *only* a problematic term. Rather, for the specific UK-based people to whom we spoke, it was also a mechanism for considering the experiences of others and for talking about political issues like in/equity. In other words, the term fairness offers a way to think and talk about data politics. This in turn raises questions about who gets to define fairness and who gets to decide what constitutes thinking and talking data politics. We conclude by proposing that participants linking matters of fairness with matters of equity could be seen as a kind of everyday data solidarity, akin to Hall's (2023) 'everyday solidarities' – that is, ordinary, unremarkable acts which might be harnessed for collective good.

Criticisms of fairness in relation to data uses

Researchers, investigative journalists and activists have demonstrated that data-driven systems disproportionately impact already disadvantaged groups (e.g. Buolamwini, 2018; D'Ignazio and Klein, 2020). When data-driven discrimination takes place in the public sector, it can be especially harmful to disadvantaged or marginalised groups,

because it can result in them being denied access to essential services, or worse. For example, in the case of the Dutch child benefits scandal 'toeslagenaffaire', benefit recipients were wrongly accused of making fraudulent claims based on data-driven risk indicators, leading to some being forced to repay benefits, some losing custody of their children and others dying by suicide (Heikkilä, 2022).

In this stark context, debates are taking place about the meaning and value of fairness. Some commentators propose ensuring fairness, accountability, transparency (FAT) – and sometimes ethics (FATE) – as a solution to data-related discrimination and harms (e.g. Mitchell et al., 2021). For example, the community formed around the ACM conference on Fairness, Accountability, and Transparency (ACM FAccT) has advanced computational methods for mitigating bias. Similarly, developers in the algorithmic fairness community have advanced technical frameworks and criteria to evaluate the fairness of algorithmic and data-driven systems (e.g. Dwork et al., 2012). Outside of FATE initiatives, fairness has been identified as a useful concept which people deploy to evaluate data uses (e.g. Kennedy et al., 2015).

The value of fairness notwithstanding, the term has been subject to criticism as both a concept and a practice. Papers such as Selbst et al.'s (2019) 'Fairness and abstraction in socio-technical systems' presents a critique of some of the assumptions driving the field of computational fairness. Like Selbst et al. (2019), other scholars have argued that approaching fairness as a computational problem that can be addressed through technical solutions fails to acknowledge the broader societal power imbalances which lead to data-driven systems having harmful outcomes (Barabas et al., 2020; Miceli et al., 2022). Indeed, Green and Hu (2018) have proposed that focusing on fairness as a technical practice limits its effectiveness; Green (2022) later argued that operating such understandings of fairness makes achieving actual fairness impossible. One of the problems here is *who* is defining fairness: when such definitions come from developers working in commercial companies, fairness can stand in tension with equity. It is for this reason that Barabas et al. (2020) and Miceli et al. (2022) propose that researchers in this field should 'study up' – that is, expand the field of enquiry to incorporate those with the power to shape data-driven systems and processes.

Criticisms of fairness as a concept also focus on its limited capacity to account for power. In conceptual terms, two criticisms are directed at fairness: first, that it is a 'floating signifier', not attached to any particular meaning and as such open to a broad range of interpretations and uses, and second, that it does not acknowledge or challenge power differentials.

The first criticism raises the question of how fairness is understood and defined. Dictionary definitions of fairness characterise it as referring to impartial and equal treatment

or behaviour, without discrimination. In these everyday definitions and related usage, fairness is linked to non-discriminatory practices. Definitions of fairness in relation to data, algorithmic and AI systems sometimes also connect fairness to equality, such as Rovatsos et al.'s (2019) Landscape Summary on Bias in Algorithmic Decision-Making, where fairness is linked to the minimisation of bias. These various definitions point to the multiple meanings of fairness. Indeed, Verma and Rubin (2018) identified more than 20 definitions of fairness in relation to data, algorithms and AI, not all of which embody a commitment to equity. It is this mutability of fairness that makes it a 'floating signifier', a term coined by Lévi-Strauss (1987) to refer to concepts 'void of meaning and thus apt to receive any meaning' (63). Unlike 'empty signifiers', which have no meaning (Laclau, 2007), floating signifiers float between meanings. Floating signifiers can mean different things depending on who is defining and mobilising them. The status of fairness as a floating signifier is seen as a problem because it opens it up to misuse. For example, writing about the terms fair and good in the context of AI, Kalluri (2020) argues that they are 'infinitely spacious' and as such, 'any AI system' can be described in these terms. In other words, a data-driven system that results in harms could still be described as fair, because it complies with one of the multiple meanings of the term.

The second criticism of fairness is that it is seen to secure rather than challenge power. In this sense, criticisms of fairness as a concept are consistent with the criticisms of fairness practices discussed above (Barabas et al., 2020; Miceli et al., 2022). Fairness is subsequently seen by some critical commentators as distinct from concepts like equity, despite the fact that fairness and equity are connected in some of the definitions summarised above. For critical commentators, the term equity is seen as a more useful conceptual tool than fairness for recognising data-related harms and acknowledging and addressing power differentials in data-driven societies (e.g. D'Ignazio and Klein, 2020; see also Costanza-Chock, 2018).

Equity itself is distinct from equality. Whereas equality means equal treatment for all, equity recognises that because of historical inequalities, resources and opportunities need to be distributed differentially in order for an equal outcome to be reached. Treating people the same can perpetuate inequalities, whereas treating people differentially can achieve 'meaningful equality of opportunity' (Menendian, 2023; see also Espinoza, 2008). Thus, equity is 'in the service of equality of opportunity'. Or, to put it another way, equality of opportunity requires equity (Menendian, 2023). We return to this distinction in the discussion of our empirical research below.

An example of this second criticism of fairness can be seen in a chart from D'Ignazio and Klein's (2020) *Data Feminism*, a specifically feminist response to computational approaches to fairness that concern other critics. In the

chart, fairness, along with accountability, transparency and ethics are included in one column under the heading 'concepts that secure power' (60). They are described as reproducing existing unequal power relationships 'because they locate the source of the problem in individuals or technical systems'. These concepts are pitted against other concepts including justice, equity and 'understanding history, culture and context', located in a separate column entitled 'concepts that challenge power' because they 'acknowledge structural power differentials and work towards dismantling them' (2020: 60). Critics argue that greater recognition of the social inequities that lead to data-related discrimination and harms is needed than 'concepts that secure power' enable. Kalluri (2020) proposes that focusing on whether data-driven systems are fair should not be the priority; rather, we need to concentrate on more political concerns like equity. Fairness is thus seen to depoliticise data and related systems and processes.

Like these critical commentators, we believe that datafication is political and that it is vital to study it within broader power dynamics. Nonetheless, we think there are some problems with the criticisms we discuss above. We question the proposition that fairness and equity are opposites, the fixed interpretations of these terms that are implied by positioning them as dichotomous, and the conclusions that are drawn based on this proposition. As noted above, definitions overlap, boundaries around concepts are blurred, and who is doing the defining matters. What's more, in our research, we found fairness to be a mechanism for many participants to think and talk politically, as we demonstrate below. We argue that defining terms like fairness from top-down, expert perspectives, as has been the case to date, overlooks the experiential, everyday understandings of members of the public that media audience researchers and sociologists of everyday life have made central to their work (e.g. Davies, 2021; Livingstone, 2018; Neal and Murji, 2015). Neal and Murji argue that social phenomena are 'made and unmade' in everyday life (2015: 812) and so too are concepts. Researching people's everyday understandings of data uses and fairness can make visible the perspectives of people from demographically disadvantaged groups, whose voices are especially excluded from relevant debates, and this can also advance understanding of perceptions of fairness and data uses. We describe how we did this below.

Methods

Our research explored perceptions of data uses in public sector organisations in the UK, focusing on people's views of specific data uses in specific domains. We did this because previous research had asked questions about data uses in abstract and general terms (Big Brother Watch & ComRes, 2015; Direct Marketing Association,

2018), which limits understanding of public perceptions of actual data uses. We focused on data uses in the government Department for Work and Pensions (DWP), the British Broadcasting Corporation (BBC) and the National Health Service (NHS), because these organisations will be familiar to many UK residents and they cover core aspects of everyday life: welfare, media and health. We partnered with the DWP and BBC to identify data in the first two domains – that is, our contacts in these organisations selected the data uses on which our research focused. For the third domain, health, we drew on publicly available information and prior research (Medina-Perea, 2021). In total, our research explored perceptions of six data uses, two each in welfare, public service media and health.

Both DWP data-based systems focused on ways of making it possible to verify identity online. The first was Confirm Your Identity, an identity verification process to enable welfare payments. The second, Dynamic Trust Hub, was a project exploring a range of ways to enhance identity verification, including technology integration and security checks. BBC data uses focused on personal control over data. The first was BBC Box, a prototype device which would pull together data about what users watch or listen to and give them control over who had access to this data to enable the generation of recommendations. The second was BBC Own It, a free app designed by the BBC to support, help and advise children when they use their phones to chat and explore the online world without adult supervision. For the NHS cases, we produced an account of the NHS Covid-19 Data Store, a national initiative storing data in one place to help coordinate the UK's Covid-19 response, by drawing on information in the public domain, on government web pages and elsewhere. We also produced an account of a data-based system in an NHS antibiotic prescribing research project which drew on prior research that one of us had undertaken (Medina-Perea, 2021).

We deployed some of the techniques used by Bates et al. (2016) as part of an approach they call Data Journeys to develop knowledge of each data use, including interviews and textual analysis of internal and publicly available documentation. We then produced written accounts and visualisations of the six data uses. We used the visualisations as elicitation tools in our empirical research and the accounts as guides for interviewers and focus group convenors to provide explanations, if required. As critical scholars of information and communications, we know it is not possible to produce objective accounts of data uses, because interpretation takes place in the act of describing, illustrating and choosing what to highlight (Bates et al., 2023). We know that methods shape findings, that 'they have effects; they make differences; they enact realities' (Law and Urry, 2004: 392–3). We are also critical of previous research which concludes that people do not mind data about them being shared and used without having

established whether participants know what happens to their data within data-driven systems and the potential harmful outcomes (e.g. Humphreys, 2011; Lupton and Michael, 2017).

Our methods are a response to these challenges. We believe that we cannot expect people to have opinions about data uses if they do not understand them, and that showing and telling people about data uses enables them to develop and express opinions about them. This is especially the case with regard to benefits and harms: people cannot assess potential benefits and harms of data uses if they do not know what they are. For this reason, we presented one proposed benefit and one potential harm of each data use to participants. For example, a proposed benefit of DWP Confirm Your Identity is that it is not necessary to confirm identity in person and with paper documents, making it easier for some people to go through this process and access the benefits to which they are entitled. A potential harm is that participation in these processes necessitates having a passport, bank account or another official document, excluding people who do not have them, perhaps because of their complex lives, making it more challenging for them to access benefits.

At the same time, we are aware that if we tell participants about the potential harms of data uses, we may lead them to feel and express concern. Providing information to our participants about data uses may have encouraged them to think about them in ways that they would not have otherwise. Some of the potential harms of DWP data uses that we introduced related to in/equity, the focus of this paper. It is important to note, then, that almost all related concerns discussed in this paper were mentioned by participants *before* we presented potential harms to them. We sought to overcome the methodological challenges we discuss here by describing and visualising data uses in ways which were clear, accurate and balanced. For example, we sourced information about benefits and harms from experts other than our partners, such as civil society data advocacy groups. We provided information for participants to reflect on while simultaneously creating space for individual and collective interpretation (see Bates et al., 2023 for extended reflection on these methodological challenges).

We began the interviews and focus groups by asking participants about their awareness of the ways in which personal data is used, their perceptions of benefits and concerns, and whether they had acted on any concerns they had. We then showed them visualisations of two or three of our public sector data uses, asked them to describe what they saw and added details as needed. We initially asked broad, open questions, such as 'What do you think of this data use?', to create space for participant-led discussion. We then introduced potential benefits and harms and asked how participants felt about the data uses in the context of this information. We concluded by asking

participants explicitly if they thought the data uses they had discussed were fair and what they thought makes a data use fair. We also asked how much they understood the data uses we discussed with them, and about the importance to them of contextual factors, such as the organisation involved, whose data is used and for what purpose. We received ethical approval from our university prior to carrying out our research (ref 032273).

Because we were interested in the role that in/equity plays in thinking about fairness and data uses, we recruited 112 demographically diverse participants. Eighty-five per cent were from one and 46% were from two or more disadvantaged groups. Often these meant participants were from groups which are minorities in the UK (for example, LGBTQ+ participants, racially minoritised participants), but not always (e.g. female participants). In summary:

- 62% participants were women, 4% were gender minorities;
- 42% were Black, Asian or from other racially minoritised groups (we presented participants with the 19 categories used in the census in England);
- 41% were on low household incomes (of less than £19,999 a year), 29% were on middle incomes (£20,000 to £59,999), and 11% were on high incomes (more than £60,000).
- 37% had a long-term condition which adversely affected their ability to carry out day-to-day activities (the 2010 UK Equality Act definition of whether someone has a disability);
- 20% were LGBTQ+ (they described their sexuality as lesbian, gay, bisexual, queer, asexual, other or answered ‘no’ to the question ‘is your gender the same as the gender you were assigned at birth?’);
- 14% were aged 65 or older.

We also gathered information about participants’ education, whether English was their first language, their job title, the ages of any children in their household, their country of birth, their sources of news and interest in politics. Sixty-six per cent of those who responded to the final two questions described themselves as interested or very interested in politics and 50% said they got their news from the BBC or *The Guardian*, relatively impartial and left-leaning sources, respectively. We asked about sources of news rather than political affiliation because the latter is a more sensitive question which may result in withdrawal from research. These figures suggest that our sample may be more left leaning than is representative of the broader UK population. Oversampling participants from minority and disadvantaged groups may have the same effects, as historically people from minoritised groups have identified with the centre-left Labour Party in the UK, although this is changing (Blayney and Evans, 2024; Turnbull-Dugarte, 2021).

Our interest in the views of people from minority and disadvantaged groups means that our sample was not

representative of the UK population. We see this as ‘understanding history, culture and context’, one of D’Ignazio and Klein’s (2020: 60) ‘Concepts that challenge power... because they acknowledge structural power differentials and work towards dismantling them’. In other words, we recruited demographically diverse participants in a particular cultural context, many of whom were interested in politics, and few of whom accessed their news from right-leaning sources. We recognise that this shaped what we found. Although our sample is not nationally representative, our participants are nonetheless members of the general public, and their views are often overlooked in debates about data uses, despite being more likely to be negatively impacted by them. It is therefore particularly important to make their opinions heard as we do here. In this sense, we suggest that our research makes a significant contribution to critical data studies, rather than seeing our sampling strategy as a limitation. Below, using pseudonyms, we share information that we gathered about participants where it is relevant to the views they expressed or to paint a picture of them. We do not share all the information we gathered to preserve anonymity, which is especially important when writing about disadvantaged or minority groups (Fox et al., 2021).

We recruited participants by posting messages on various social media groups and pages (such as groups for Latin Americans in the UK and LGBTQ+ groups), contacting charities and organisations which worked with groups of people we wanted to reach and leafleting targeted homes, such as council-owned tower blocks, once lockdown restrictions were eased. We made provisions to carry out our research, undertaken online because of Covid-19, in ways which were inclusive of participants with limited access to devices and the internet. These included phone interviews, posting print-outs of visualisations, setting up a research mobile phone number for participants who wished to communicate this way and working with facilitators to enable people with cognitive or learning disabilities to participate. Once pandemic conditions allowed, two face-to-face focus groups were conducted with women on low incomes who otherwise would not have been able to participate.

We carried out thematic analysis of the data we gathered in NVivo, developing codes that were grouped into themes deductively and inductively. We approached our thematic analysis with an intersectional lens, following Crenshaw (1991) and others who acknowledge that systems of power and inequity – for example, relating to gender, race and ethnicity, and class – intersect to shape experiences of discrimination and privilege. We looked at how inequities intersected in particular cases to support our thematic analysis, an approach which enabled us to meet our overarching aim of exploring how diverse members of the UK public perceived specific public sector data uses and what fair data uses look like from their perspectives. Some

intersections stood out in our analysis as affecting views of fairness, and we highlight these in our discussion of findings below.

Connecting fairness and equity in perceptions of data uses

In this section, we first describe the ways in which in/equity surfaced and mattered in participants' reflections on the public sector data uses that we discussed with them. We then demonstrate how, when asked their thoughts about what constitute fair data uses, the term fairness had multiple meanings. These multiple meanings made it possible for many participants to link fairness to concerns about equity. In the third section, we show how participants made these connections. Sometimes they did this in response to questions we asked about fairness, but often they did it spontaneously, without us asking.

Experiences of in/equity and perceptions of data uses

Belonging to one or more structurally disadvantaged or minority group appeared to inform many participants' reflections on data uses. For example, some older participants and participants with long-term conditions were more comfortable with health data uses than data uses in other sectors, perhaps because these participants depended on health services. In contrast, health data uses concerned LGBTQ+ participants more than heterosexual cisgender participants. Some Black, Asian and other racially minoritised participants were concerned about the ways in which structural racism is reproduced through data uses. Some participants with long-term conditions, or for whom English was an additional language, or who had low levels of education, indicated that these aspects of their identities made understanding and engaging with data uses challenging.

For example, Gulay is a Turkish-born heterosexual woman, aged 35–44, a student on a low income with long-term conditions. She noted that she found understanding data uses hard, particularly as they are described in terms and conditions. At the same time, knowing that data about her was secure was important to her because of her status as a refugee. Talking about companies' online privacy policies, she said:

This is hard for me because it's in English and long. I know I am not accept, not continue. I just accept. [...] I'm refugee. I come to this country because of problems in my country. I don't want my location to be shared with this [Turkish] government. This is important for me. I know this country is very safe and – it is okay. [...] But I don't want to share with other countries, like my country.

Gulay's status as a refugee intersected with the fact that English was an additional language for her, something

which confounded the challenge of making sense of complex information about data uses. These intersections combined with her biographical experience in a less 'safe' country, leading her to express concern about the safety of data about her. Gulay's words point to the way in which intersecting aspects of her identity (Crenshaw, 1991) informed her perceptions of data uses.

Other participants also recognised the importance of identity intersections in experiences and perceptions of data uses. Louisa, a white, heterosexual woman who was born in the UK, aged 35–44, who works for a charity, is on a middle income and has no long-term conditions, acknowledged the challenge of making sense of already complex data uses for people for whom English is an additional language, who may be refugees, and whose lives may already be complex because they are homeless or in need of welfare support. Heidi, a queer, white British woman born in the UK, aged 25–34, who works as a doctor, has a high income and no long-term conditions, reflected on the data-related vulnerabilities that emerge at the intersection of LGBTQ+ identity and other factors. Heidi recognised that her experience as a 'financially and socially comfortable' queer woman would not be shared by all queer women:

I imagine still within the UK there is definitely a subset of people where, actually, their community wouldn't accept them being gay and actually [...] if they are then Googling gay websites or something and then that flips into their advertising and their family member picks up their phone, actually that is then really important.

Heidi also noted the data-related challenges that younger queer women might experience, whose queer identities might be revealed to their parents through data processes:

I think that's something that can be potentially really harmful because those parents can react in a bad way. You can get abuse, children that are kicked out from their homes, from data that's shared back.

Likewise Tahira, a heterosexual, Pakistani woman, aged 45–54, with a high household income, reflected on the extra security checks that were being considered as part of DWP Dynamic Trust Hub (shown in Figure 1), which included checking time of log in, the rhythm with which people type their passwords, swiping patterns and the devices that are used to log in. Although not on a low income herself, she was concerned about how difficult it would be for people on low incomes with multiple jobs to meet the requirements of these checks.

Heidi and Tahira recognised their advantages with regard to income and employment. They felt empathy for those for whom this is not the case and who thus might experience data uses more negatively. We also found that

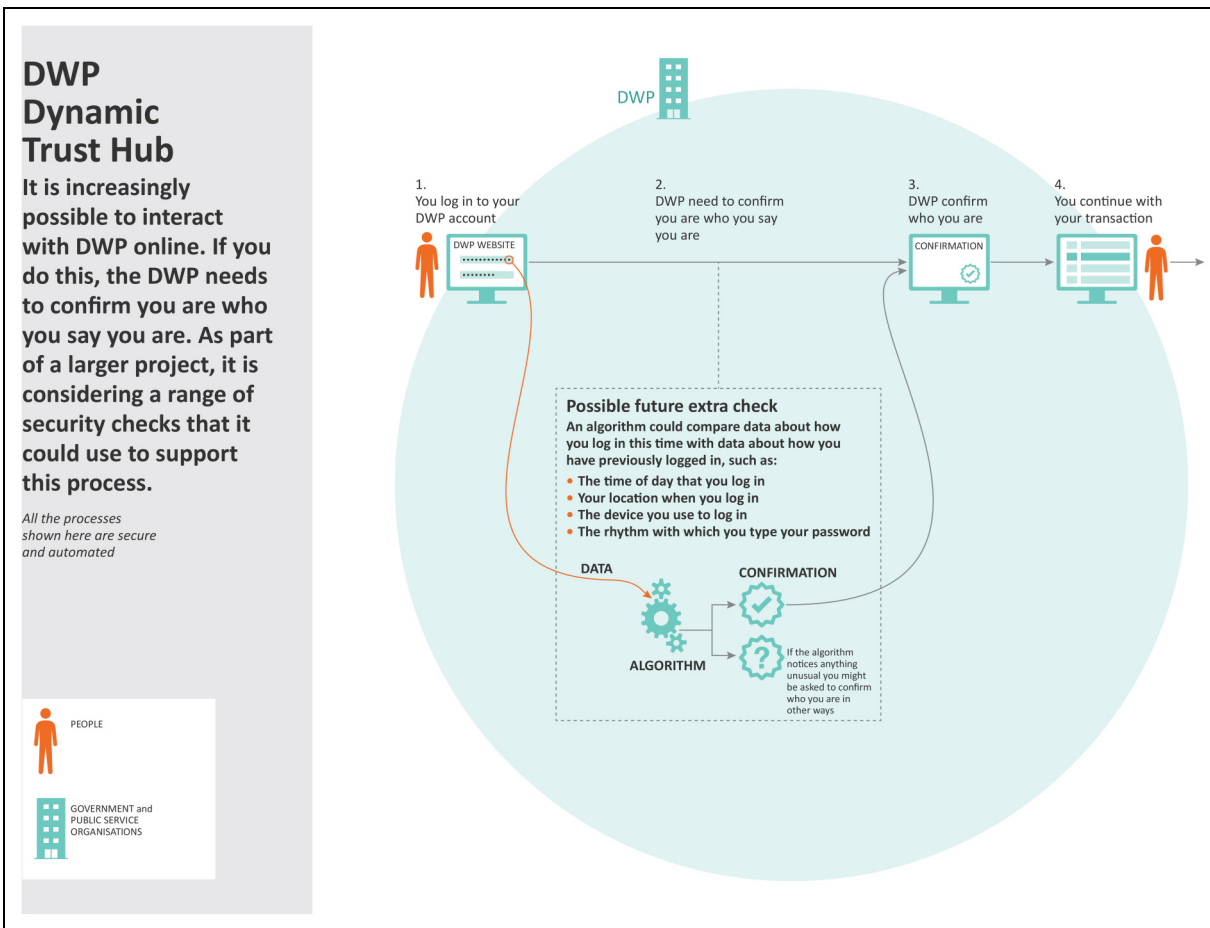


Figure 1. DWP Dynamic Trust Hub, a project exploring ways to enhance identity verification.

participants from one disadvantaged or minority group were sometimes concerned about the effects of data uses on people from another disadvantaged or minority group. For example, Huso, a Black British African, heterosexual woman, aged 25–34, who works as a teacher and has no long-term conditions, expressed concern about people with English as an additional language and limited technical resources attempting to use DWP’s data-driven systems.

Some participants felt that their own privilege gave them an advantage in engaging with data processes which other people did not have, like Heidi who is quoted above. Rosie, a white British heterosexual woman, aged 35–44, a physiotherapist with a middle income and no long-term conditions, described how complicated she and her husband found the online system for accessing government services which it is necessary to log into as part of DWP Confirm Your Identity:

It took us forever and we’re like two, we’ve both got degrees and have been through university and English is our first language. I dread to think for somebody who isn’t, you know, who hasn’t got English as a first language. It’d just be horrendous.

Rosie did not belong to a disadvantaged or minority group but she was nonetheless concerned about how people from these groups might be negatively impacted by data uses. Likewise Craig, a heterosexual, White British man, aged 35–44, with a middle income, an undergraduate degree and no long-term conditions, expressed concern about people with long-term conditions being discriminated against as a result of certain health data uses, for example if such data was shared with health insurance companies. Fran, a White British, heterosexual woman, aged 35–44, with a BTEC diploma and a high income, was concerned about whether people on low incomes could engage with DWP data processes that were essential to access welfare support. She said: ‘if you’re in financial difficulties and you don’t have the money to top up your phone or to have internet access, it’s making it very hard to get on and get the support you need’.

Participants were particularly concerned about in/equity in relation to DWP data uses, because DWP provides essential welfare services to the very people who are likely to be disadvantaged by structural inequities.

In/equities mattered in participants’ reflections on data uses in various ways. For example, belonging to a

structurally disadvantaged group whose members are more likely to experience the harmful effects of data uses informed perceptions, as in Gulay's case. We found concern about the negative consequences of some data uses for people from disadvantaged or minority groups amongst participants who were not from those groups. This is an effect of who our participants were: for some, their own disadvantage made them aware of other disadvantages. However, participants who were not from disadvantaged groups – Louisa, Rosie, Craig and Fran – were also concerned about how these groups might be more negatively impacted by data uses than others. This confirms a finding from our quantitative research which we report elsewhere (Taylor et al., 2023). Awareness of inequities, as well as experiences of them, led to concern that particular data uses might reproduce them. Below, we argue that these concerns connected to participants' perspectives on fairness and that the multiple meanings of fairness made these connections possible.

Fairness as floating signifier: The multiple meanings of 'fair' data uses

When we asked whether they thought that the public sector data uses we discussed with them were fair, a small number of participants were confused about how fairness could connect to data processes. Melissa, a white bisexual woman who was born in the USA, aged 35–44, a psychotherapist with a middle income and no long-term conditions, asked 'Fair in what way, how do you mean?' Diane, a white, heterosexual woman who was born in the UK, aged 55–64, on a middle income and with no long-term conditions, said 'It depends on how you define fair, doesn't it, really?' It is because fairness is a floating signifier that these participants were not sure what we meant by it. Jim, a heterosexual man aged 65+ who identified as European, on a middle income and with no long-term conditions, also expressed uncertainty about how fairness could be applied to the context of data uses:

I'm not sure what you mean by fairness in this context. I'm not quite sure what – how the word fairness fits into this. I'm not entirely sure what you mean. I mean, life's not fair, is it, you know. I mean, some people are born rich, some people are born disabled, some people are born poor, you know. I don't know what you mean by fairness in this context.

However, in the conversation that followed, as Jim's friends discussed what constitutes a fair data use, Jim joined in, saying that he was concerned about unscrupulous people getting access to data, which suggests that, for him, using data responsibly was a component of fair data uses. He had a view on what makes a data use fair, even though he struggled to see the relevance of the term fairness in this context. This was also the case for Melissa and

Diane, as Melissa went on to describe choice as a component of fairness and Diane linked fair data uses to transparency. Thus, the small minority of participants who questioned the concept of a fair data use still had something thoughtful to say about it.

Most of our participants did not struggle with our question about fair data uses and gave direct answers. Mbali, a Black, British African, heterosexual woman, aged 25–34, a student with a middle income and no long-term conditions, noted that the purpose for which data is gathered is an important factor:

I think it's the purpose of the use. We live in capitalist societies, so somebody is always after your money in some way or form [laughs], and in that way, I think commercially it's very difficult for your data to be used fairly, 'cos it is always in some way manipulative. But in terms of the public sector, I think, if they are not led by profit and they're led by trying to provide certain services, I think that changes it. [...] I guess it's just about what the motivation is behind them doing certain things, and if that's fair, then the whole process should be fair.

Tanya, a white, lesbian woman who was born in the UK, aged 55–64, who works in the public sector, has a long-term condition and a middle income, responded to the same question by describing a fair data use as 'for the greater good [...] for health and benefits, for people, for us to learn, develop and grow in a way that's going to make society a fairer place'. Some participants said clear information which makes it possible to understand what happens to data was a component of a fair data use. Lewis, a white, heterosexual man who was born in the UK, aged 55–64, who had a high income and no long-term conditions, felt that the characteristics of fair data uses included "transparency and honesty in terms of what's been captured, how long it's been kept for, and what purpose it's going to be used for". For Lewis, being able to choose *both* what happens to your data *and* to share data once you understand how it will be used were important components of fairness.

Kahina, a Black, heterosexual woman who was born in Somalia, aged 25–34, a student on a middle income and with no long-term conditions, compared the two BBC data uses which experimented with personal control over data (Figures 2 and 3) with the DWP Dynamic Trust Hub (Figure 1), which focused on ways to verify identity online of choice. With BBC Box and BBC Own It, people 'have got a bit of a choice' with regard to whether they complete or download a profile, she said. Because engagement with DWP "is a need" for many people, and potential Dynamic Trust Hub security checks may not be optional, there is less choice, which, in Kahina's eyes, made it more unfair.

For many participants, fairness was not just one thing – that is, not just clear information, being given choices,

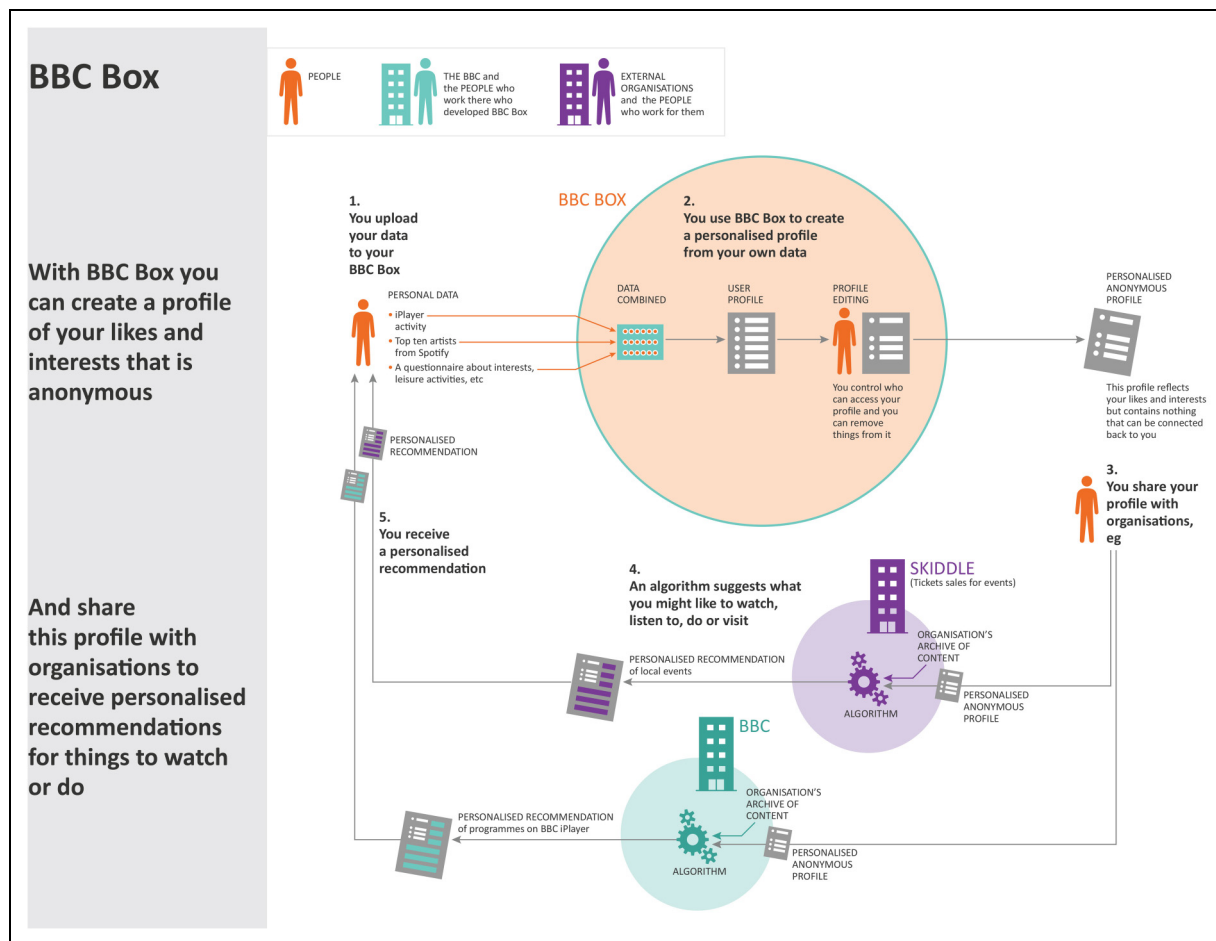


Figure 2. BBC Box, which pulls together data about what users watch or listen to and gives them control over who has access to this data.

being for the social good, nor being inclusive. In most responses to questions about what makes data uses fair, participants discussed more than one of fairness' multiple meanings. Louisa, mentioned above, said that consent, choice and understanding what is being consented to were equally important characteristics of fairness. Ahmed, a Pakistani heterosexual man who was born in the UK, aged 25–34, unemployed and looking for work, with a low income and no long-term conditions, said that fairness involves clear information, transparency about what data will be used for, data being used for the social good and the ability to consent:

It's where people are kept well informed and data is used for a valuable or useful purpose, and there's a transparent process as well for sharing the data. And obviously permissions are sought too.

As can be seen, there is more than one understanding of fairness. It is not an empty signifier, attached to no meaning at all (Laclau, 2007). Rather, it is a floating signifier, with

different meanings for different people. It is not 'infinitely spacious' as Kalluri (2020) claims, as certain characteristics of a fair data use were repeated: clarity of information, offering choice with regard to consent, for the social good. The openness with regard to what fairness signifies made it possible for participants to link fairness to concerns about equity, as we demonstrate below. In the quotes above, we can already see these linkages emerging, such as when Jim talked about inequities in income and ability as constitutive of an unfair society, and when Tanya connected fair data uses to broader societal issues such as the social good.

Linking fairness and inequity in talk about data uses

Because of the multiple meanings of fairness, some of our participants linked it to equity in their reflections on the data uses we discussed with them. When we asked them what makes a fair data use, some participants spoke about the need for data uses to work 'for all communities', sometimes listing groups who they perceived to be excluded from certain data processes, as seen above. This shows

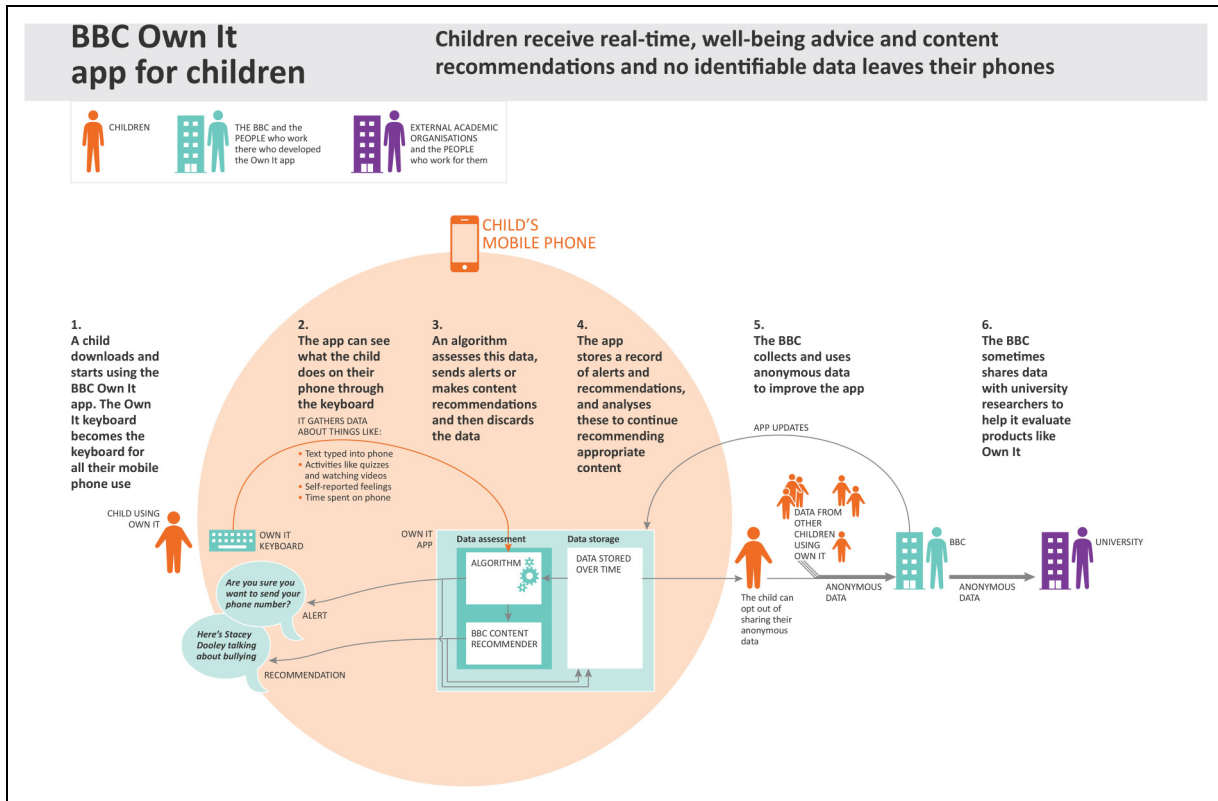


Figure 3. BBC Own It, a free app designed to support children when they use their phones, without adult supervision.

that thinking about fairness involved thinking about historical inequities, for some participants. These participants appeared to be aware that data uses can reinforce or deepen inequities, for example related to income, education or ability, and that some data uses, such as those in the welfare sector, are more likely than others to do so, and they referred to these concerns when we asked their views about what constitutes a fair data use. Most participants did not explicitly use the terms equity or inequity. Nor would we expect them to, as these terms might represent ‘a rather alienated way of describing things’ (Sayer, 2011: 2) to people for whom such language may not be part of everyday vocabulary. Rather, we argue that there are connections between fairness and equity in what participants said when responding to questions about what fair data uses look like.

For example, when asked whether they thought the data processes we discussed with them were fair, participants often considered how specific data uses might impact people from disadvantaged groups and concluded that they were unfair if these groups could be disproportionately affected. Tanya, mentioned above, saw data processes which are not equally accessible to all as unfair. Tanya described the NHS antibiotic prescribing research project, which explored ways to address antibiotic resistance (Figure 4), as fair because of her ‘sense of where it’s

motivated from’ – that is, it used data for what she perceived to be the public good. In contrast, she felt that DWP Confirm Your Identity, an online identity verification process for welfare payments (Figure 5), was not ‘completely fair and all-inclusive for all communities’, because it required technology access and English language skills that not all people have. She said:

People don’t have access to technology in the first place. And you know, what about people with English as an additional language or no English, or families with additional needs, where using technology might be difficult, or they aren’t literate.

Tanya made this comment after we had pointed out that some groups were concerned that identity checks are not accessible to all, such as people without credit histories, with unusual residence histories or otherwise complicated lives. But it was Tanya who linked the notion of a fair data use to issues of inequity in what she said about Confirm Your Identity.

In another example, Rosie, mentioned above, participated in a discussion about the two BBC examples and DWP Dynamic Trust Hub (Figures 2, 3 and 1, respectively). When asked what she thought would make these data uses fair, she said that information about them

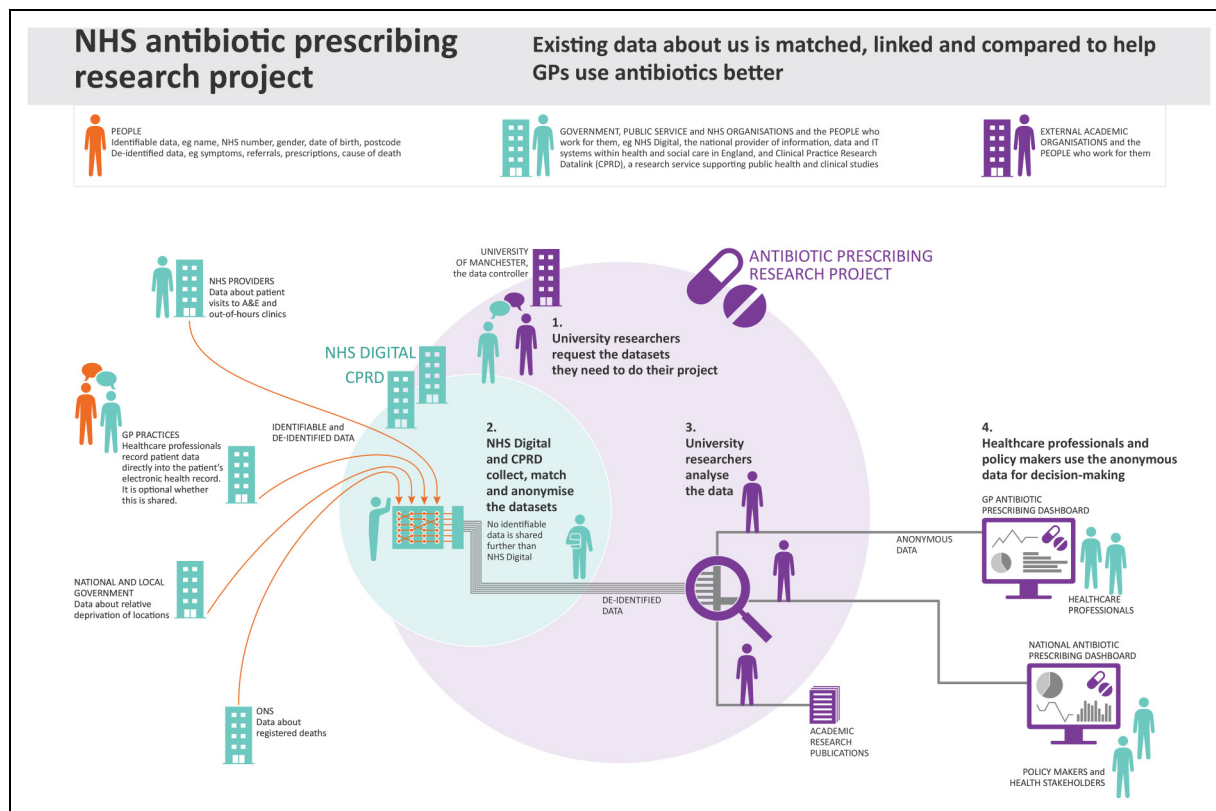


Figure 4. NHS antibiotic prescribing research project, exploring ways to address the public health crisis of antibiotic resistance.

needed to be available in different languages in order to give people ‘fair and equal access to understanding what they’re sharing’. For Kerry, a white, British, queer woman aged 35–44, with a high household income, who worked in the charity sector and had no long-term health conditions, the welfare context of Dynamic Trust Hub was significant. Because already disadvantaged or vulnerable people are dependent on welfare for their survival, it is especially important that data uses in this context are accessible, she felt. Grace, a white, British, bisexual woman, aged 35–44, a teacher with a high household income and no long-term conditions, thought that the DWP Confirm Your Identity data use was unfair, because ‘not everyone can access it’. Teddy, a white British, heterosexual man, 65+, retired and on a high income, also linked the fairness of data processes to people’s ability to access them. In response to the question ‘are any of the data uses we have talked about fair?’, Teddy noted that ‘The benefit of the BBC’s Own It in telling me what to do if I’m bullied is only available to people who have phones. So kids who don’t have phones are disadvantaged.’ In various ways, these participants link their reflections about fairness to in/equity, especially with regard to access to data-driven systems.

The participants cited above responded to questions we asked them about what would constitute a fair data use by

bringing up concerns about equitable access. Other participants used the words fair and unfair without being prompted to do so by us, connecting these terms to the inequities that they saw in the data uses in similar ways. For example, Ruby, a heterosexual, British Chinese woman, aged 18–24, who works in legal services, with a middle income and no long-term conditions, expressed concern about the implications of DWP Confirm Your Identity for people who do not have documentation such as a passport or tax records. She said:

those people that don’t have a P60 [a statement issued to UK taxpayers at the end of the tax year], is it more people that have jobs that are cash in hand, for example, and people that don’t have a passport, would they, you know, have never left the country or never needed a passport. That makes me think of people that may be [...] if you’ve never been abroad and have a cash in hand kind of job, that makes me think the people that are closer to the poverty line have to part with their data more than people that have like a P60, passport, like you said, and that’s really unfair.

Ruby’s belief that people without the documentation to enable online identity verification would have to ‘part with their data more’ might not be entirely accurate, but

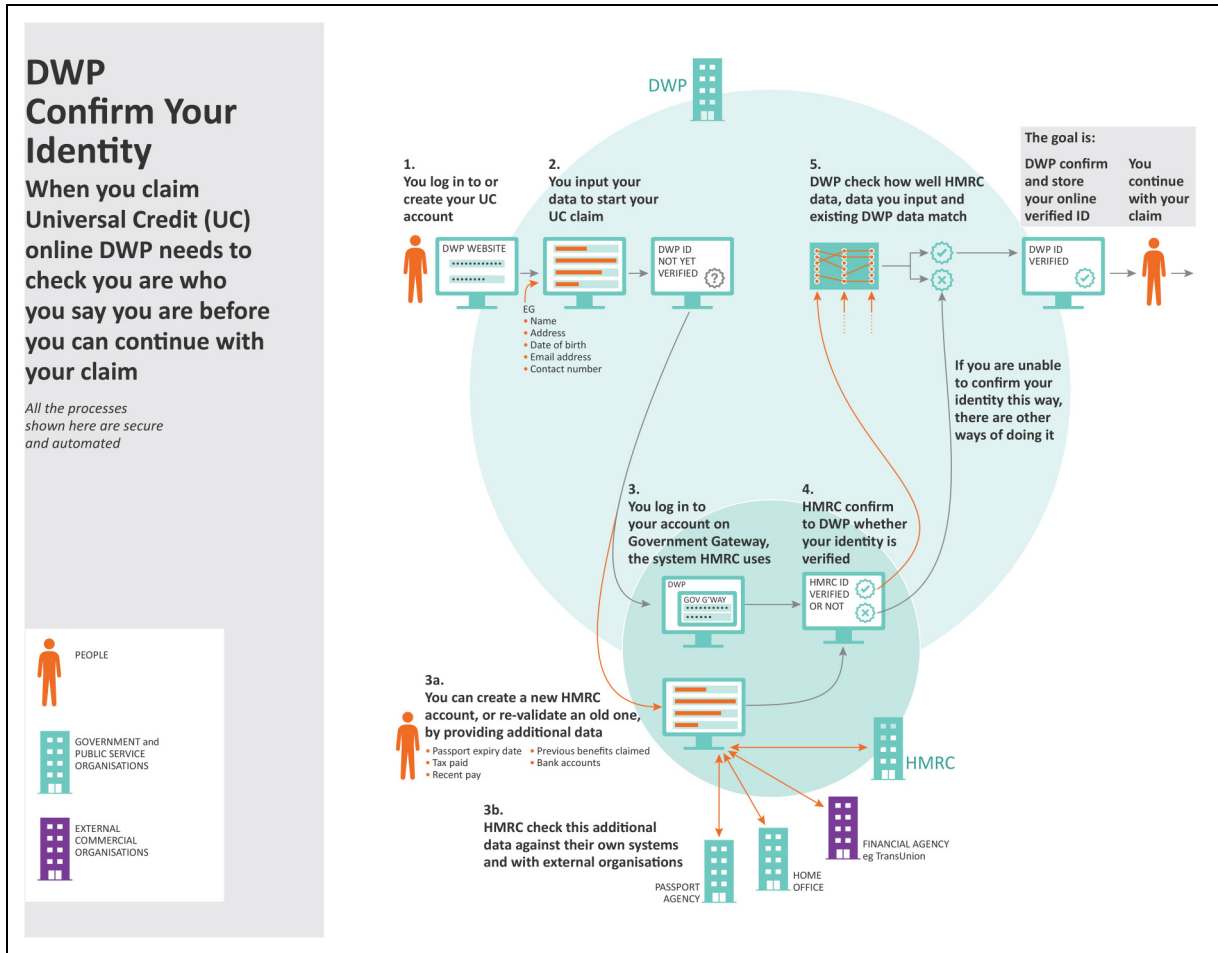


Figure 5. DWP Confirm Your Identity, an identity verification process for welfare payments which makes it possible to confirm identity online.

what is important here is that she links structural inequities to unfairness in data processes. Tahira, mentioned above, did the same. She was positive about BBC Own It, an app designed to support children when they use their phones to explore the online world, because she felt that it might overcome inequities by providing access to resources to people who otherwise might not be aware of or able to access them:

A lot of people I know wouldn't normally have access to that kind of resource. [Pakistani parents sitting at home in the UK] wouldn't know where to reach out to, because they've not been educated in this country, for example, or just don't know. [...] So, for me that's the fair one, if I was to look at it from that lens. [...]

In the examples above, participants linked fairness and in/equity, either in response to questions we asked, or voluntarily while reflecting on inequities. We argue that it is equity, not equality, that concerns participants, because they recognise that historical inequities influence people's

ability to engage in data processes, and that these structural factors need to be addressed (Menendian, 2023). Whereas critical data study scholars propose that the concepts of fairness and equity are distinct, these terms came together in the comments of many of our participants. In contrast to Kalluri's (2020) proposition that asking whether data and AI processes and systems are fair is the wrong question, we found that asking this question opened up a space for participants to consider the politics of datafication that concern critical commentators. Therefore, for many of our participants, talking about fairness meant talking about data politics. Most of the participants cited above described themselves as very or fairly interested in politics; only Gulay and Diane did not. They accessed their news from a range of sources, including the BBC, *The Guardian* and other news websites and outlets. It is not surprising, therefore, that they talked data politics in our conversations with them.

Furthermore, *how* they talked data politics is important. The participants we cite above did not use a politicised or social scientific vocabulary – of inequity or power, for

example – but nonetheless, thinking about whether data processes are fair or not involved thinking politically, thinking about disadvantaged groups or about collective concerns. Our participants used or responded to the concept of fairness in ways that challenge assumptions about what being concerned about data politics looks like. The term fairness may fail to acknowledge structural power in some technical mobilisations (Barabas et al., 2020; Miceli et al., 2022), but it can also express concern about the ways in which data uses relate to structural inequities in others. Dismissing the concept of fairness, as some critics do, runs the risk of dismissing people’s political concerns. Below, we suggest that connecting fairness and equity could be seen as a form of everyday data solidarity.

From fairness and equity connections to everyday data solidarity

In this paper, we advance understanding of fairness in relation to data systems and processes. Fairness in relation to data uses has been criticised because it is a floating signifier, not attached to any particular meaning and as such open to a broad range of uses, and because it is deployed in ways which fail to acknowledge structural power, thus depoliticising datafication. Critical data study scholars distinguish fairness, understood as ‘locating the source of the problem in individuals or technical systems’, from concepts like equity and justice, seen to ‘acknowledge structural power differentials and work towards dismantling them’ (D’Ignazio and Klein, 2020: 60).

For the demographically diverse group of participants to whom we spoke about public sector data uses, this distinction did not hold true. We found that participants responded to our questions about fair data uses by talking about social inequities, which challenges Kalluri’s (2020) argument that asking whether data and AI processes and systems are fair is the wrong question. In fact, we found that asking these questions created space for some participants to discuss the politics of datafication. At the same time, some participants introduced the terms fair and unfair to the conversation, using them to describe what they saw as inequitable data uses, without being prompted to do so by us.

Thus, fairness and equity were not so distinct in our participants’ perceptions of public sector data uses as they are seen to be in the work of critical data study scholars. We note above that the two concepts are linked to each other in a range of definitions, and that the boundaries between the two terms are somewhat blurred. We argue that one of the reasons for this is because fairness is a floating signifier. Its status as such is therefore not only problematic, it is also potentially productive. We have shown that the terms fair and unfair were used to express concern about inequities, sometimes when prompted by us, at other times voluntarily. We further suggest that participants linking matters of fairness with matters of equity could be seen as a kind

of everyday data solidarity. This is not the active, activist solidarity that concerns many commentators, but a more everyday kind, akin to what Nikunen (2019) calls social solidarities – that is, not experienced consciously as *solidarities*, and as such distinct from more conscious and active political manifestations.

There are various characteristics of solidarity that lead us to describe what we found in our research as everyday data solidarities. For example, solidarities are said to emerge across similarities and differences. Prainsack and Buyx (2016) define solidarity as a practice that reflects people’s commitments to supporting others with whom they recognise similarity, but Littler and Rottenberg argue that ‘to express solidarity with others is ostensibly to recognize and respect differences without colonizing those differences’ (2020: 865). In our research, we saw both: participants cared about how data uses might impact people from other demographic groups as well as from their own.

Although the commonplace association of solidarity with political action (e.g. Prainsack et al., 2022) may mean it seems ill-fitting in the context of our research, solidarity is also said to incorporate the *potential* for action (Littler and Rottenberg, 2020). Conceiving of friendships as everyday solidarities, Hall (2023) argues that they hold within them everyday political possibilities, of ordinary, unremarkable acts which might be harnessed for collective good. She draws on Hankins (2017) to describe these as the ‘quiet politics’ of everyday life. In short, as Littler and Rottenberg argue, solidarity – like fairness – is a plural and ‘capacious’ concept, which we suggest could be productively applied to our research findings.

At the same time, it is important to acknowledge the potential limitations of the concept of solidarities. Understanding oneself as better informed than others, as our participants Heidi and Rosie did, or expressing concern about the consequences of data uses for others from disadvantaged groups, as many participants did, may risk becoming somewhat paternalistic. Scholars of care – which shares a number of the characteristics of everyday solidarities – acknowledge that care privileges some and excludes others, noting ‘its lack of innocence and the violence committed in its name’ (Martin et al., 2015: 627). In contrast, the limitations of solidarity are rarely recognised (Highfield and Miltner, 2023 is an exception).

Oman and Bull (2022) describe identifying connections, as we did in our research, as ‘standing in the gap’, or focusing on what connects phenomena rather than what separates them. They argue that to stand in the gap is ‘to act as a mediator whilst also standing in defence of someone or something’ (2021: 34). Taking such an approach accounts for lived experience whilst simultaneously joining phenomena together, rather than setting them in opposition to each other, they claim. Following Oman and Bull, it could be argued that focusing on the connections between fairness and equity in how people reflect on data uses could be more productive than describing these concepts as opposites when it comes

to securing or challenging data power. Given the productive potential of talking about fairness that we demonstrate in this paper, we argue that in fully dismissing the term, there is a danger that what matters to people about data uses is also dismissed. In our research, we found that what mattered to our participants was often a concern about inequities. When these concerns are expressed by people from the structurally disadvantaged and minoritised groups which are most impacted by data-related harms, dismissing them is especially troubling – doing so could leave fewer conceptual tools with which to talk to people about data politics than would otherwise be the case. What is at stake here is who gets to define what fairness means – when defined by diverse members of the public who are not data experts, it has productive, political meaning.

Because we found that the term fairness offered a way for some of our participants to talk about data politics, our research also raises important questions about who gets to decide what counts as thinking and talking about data politics. Critical data study scholars have an interest in understanding and intervening in data power, and thinking with *both* fairness and equity might enable greater understanding of the perspectives of the people at the heart of datafied societies.

Acknowledgments

The authors would like to thank Living with Data team member Lulu Pinney for producing the visualisations used in this research.

Data availability statement

Data are available through ORDA, the University of Sheffield's data repository, at doi: 10.15131/shef.data.c.6122043.v3.








Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Nuffield Foundation (grant number OSP/43959).

ORCID iDs

Helen Kennedy  <https://orcid.org/0000-0003-0273-3825>
 Hannah Ditchfield  <https://orcid.org/0000-0001-8612-620X>
 Susan Oman  <https://orcid.org/0000-0002-5723-0830>
 Jo Bates  <https://orcid.org/0000-0001-7266-8470>
 Itzelle Medina Perea  <https://orcid.org/0000-0003-1702-1484>
 Monika Fratzak  <https://orcid.org/0000-0001-6156-0982>
 Mark Taylor  <https://orcid.org/0000-0001-5943-9796>

References

Barabas C, Doyle C, Rubinovitz JB, et al. (2020) Studying up: Reorienting the study of algorithmic fairness around issues of

- power. In: ACM Conference on Fairness, Accountability, and Transparency, January, pp.27–30.
- Bates J, Kennedy H, Medina Perea I, et al. (2023) Socially meaningful transparency in data-based systems: Reflections and proposals from practice. *Journal of Documentation* 80(1): 54–72.
- Bates J, Lin Y-W and Goodale P (2016) Data journeys. *Big Data and Society* 3(2). <https://doi.org/10.1177/2053951716654502>.
- Big Brother Watch & ComRes (2015) UK public research—Online privacy. <https://www.bigbrotherwatch.org.uk/wp-content/uploads/2015/03/Big-Brother-Watch-Polling-Results.pdf>
- Blayney M and Evans G (2024) The state of public opinion: Class. *UK in a changing Europe*. <https://ukandeu.ac.uk/the-state-of-public-opinion-class/>
- Buolamwini J (2018) Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research* 81: 1–15.
- Costanza-Chock S (2018) *Design Justice: Community-Led Practices to Build the Worlds We Need*. Boston: MIT Press.
- Crenshaw KW (1991) Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review* 43(6): 1241–1299.
- Davies K (2021) Sticking together in ‘divided Britain’: Talking Brexit in everyday family relationships. *Sociology* 56(1): 97–113.
- Dencik L, Hintz A, Redden J, et al. (2019) Exploring data justice: Conceptions, directions, applications. *Information, Communication and Society* 22(7): 873–881.
- D’Ignazio C and Klein L (2020) *Data Feminism*. Boston: MIT Press.
- Direct Marketing Association (2018) *Data privacy: What the consumer really thinks*. https://dma.org.uk/uploads/misc/5a857c4fdf846-data-privacy—what-the-consumer-really-thinks-final_5a857c4fdf799.pdf
- Dwork C, Hardt T, Pitassi T, et al. (2012) Fairness through awareness. In: Proceedings of the 3rd innovations in theoretical computer science conference, pp.214–226: ACM.
- Espinoza O (2008) Solving the equity-equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research* 49(4): 343–363.
- European Union (2016) General data protection regulation. *Official Journal of the European Union* 49: L119.
- Fox J, Pearce K, Massanari A, et al. (2021) Open science, closed doors? Countering marginalization through an agenda for ethical, inclusive research in communication. *Journal of Communication* 71(5): 764–784. doi:10.1093/joc/jqab029.
- Green B (2022) Escaping the impossibility of fairness: From formal to substantive algorithmic fairness. *Philosophy and Technology* 35(90). doi:10.1007/s13347-022-00584-6.
- Green B and Hu L (2018) The myth in the methodology. In: Machine Learning: The Debates Workshop, 35th International Conference on Machine Learning. https://econcs.seas.harvard.edu/files/econcs/files/green_icml18.pdf.
- Hall SM (2023) Friendship, academic and everyday solidarities: Friendship as feminist praxis. Unpublished talk.
- Hankins K (2017) Creative democracy and the quiet politics of the everyday. *Urban Geography* 38(4): 502–506.
- Heikkilä M (2022) Dutch scandal serves as a warning for Europe over risks of using algorithms. <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>.

- Highfield T and Milner K (2023) Platformed solidarity: Examining the performative of Twitter hashtags. *Convergence* 29(6): 1641–1667.
- Humphreys L (2011) Who's watching whom? A study of interactive technology and surveillance. *Journal of Communication* 61(4): 575–595.
- Kalluri P (2020) Don't ask if artificial intelligence is good or fair, ask if it shifts power. *Nature* 583(7815): 169. doi:10.1038/d41586-020-02003-2.
- Kennedy H, Elgesem D and Miguel C (2015) On fairness: User perspectives on social media data mining. *Convergence* 23(3): 270–288.
- Laclau E (2007) *Emancipation(s)*. London: Verso.
- Law J and Urry J (2011) Enacting the social. *Economy and Society* 33(3): 390–410. doi:10.1080/0308514042000225716.
- Lévi-Strauss C (1987) [1950] *Introduction to Marcel Mauss*. London: Routledge.
- Littler J and Rottenberg C (2020) Feminist solidarities: Theoretical and practical complexities. *Gender, Work & Organisation* 28(3): 864–877.
- Livingstone S (2018) New media, new audiences? *New Media & Society* 1(1): 59–66.
- Lupton D and Michael M (2017) 'Depends on who's got the data': Public understandings of personal digital dataveillance. *Surveillance and Society* 15(2): 254–268.
- Martin A, Myers N and Viseu A (2015) The politics of care in technoscience. *Social Studies of Science* 45(5): 625–641.
- Medina-Perea I, 2021. *Socio-material factors shaping patient data journeys in the United Kingdom*. PhD Thesis. University of Sheffield. <https://etheses.whiterose.ac.uk/30111/>
- Menendian S (2023) Equity vs equality: What's the difference? Othering & Belonging Institute blogpost. <https://belonging.berkeley.edu/equity-vs-equality-whats-difference>.
- Miceli M, Posada J and Yang T (2022) Studying up machine learning data: Why talk about bias when we mean power? *Proceedings of ACM Human-Computer Interaction* 6(34): 1–14.
- Mitchell S, Potash E, Barocas S, et al. (2021) Algorithmic fairness: Choices, assumptions, and definitions. *Annual Review of Statistics and Its Application* 8: 141–163.
- Neal S and Murji K (2015) Sociologies of everyday life: Editors' introduction to the special issue. *Sociology* 49(5): 811–819.
- Nikunen K (2019) *Media Solidarities: Emotions, Power and Justice in the Digital Age*. London: Sage Publications.
- Oman S and Bull A (2022) Joining up well-being and sexual misconduct data and policy in HE: 'To stand in the gap' as a feminist approach. *The Sociological Review* 70(1): 21–38.
- Patel R and Jones E (2021) *The Data Divide*. London: Ada Lovelace Institute. <https://www.adalovelaceinstitute.org/report/the-data-divide/>.
- Prainsack B and Buyx A (2016) Thinking ethical and regulatory frameworks in medicine from the perspective of solidarity on both sides of the Atlantic. *Theoretical Medicine and Bioethics* 37(6): 489–501.
- Prainsack B, El-Sayed S, Forgó N, et al. (2022) Data solidarity. White Paper. <https://www.governinghealthfutures2030.org/wp-content/uploads/2022/12/DataSolidarity.pdf>
- Rovatsos M, Mittelstadt B and Koene A (2019) Landscape summary: Bias in algorithmic decision-making. <https://www.research.ed.ac.uk/en/publications/landscape-summary-bias-in-algorithmic-decision-making-what-is-bia>.
- Sayer A (2011) *Why Things Matter to People: Social Science, Values and Ethical Life*. Cambridge: Cambridge University Press.
- Selbst AD, Boyd D, Friedler SA, et al. (2019) Fairness and abstraction in sociotechnical systems. In: Proceedings of the conference on fairness, accountability, and transparency, pp.59–68.
- Taylor L (2017) What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society* 4(2). <https://doi.org/10.1177/2053951717736>.
- Taylor M, Kennedy H and Oman S (2023) Challenging assumptions about the relationship between awareness of and attitudes to data uses amongst the UK public. *The Information Society* 40(1): 32–53.
- Turnbull-Dugarte SJ (2021) The UK's sexuality gap: LGB voters remain significantly more supportive of the Labour Party. *The LSE Impact Blog*. <https://blogs.lse.ac.uk/politicsandpolicy/uk-lgb-voters/>.
- Verma S and Rubin J (2018) Fairness definitions explained. In: *FairWare'18: IEEE/ACM International Workshop on Software Fairness*. Sweden: Gothenburg.