



Context and prior awareness matter in public concerns about data uses

Findings from the 2020 Living With Data Survey

Helen Kennedy, Mark Taylor, Susan Oman, Jo Bates, Itzelle Medina-Perea,
Hannah Ditchfield, Lulu Pinney

June 2021

KEY FINDINGS

- 1. Public concern about certain data uses remains high.** Survey respondents are highly concerned about particular aspects of data use and data uses in particular contexts. High levels of concern have also been identified in previous research. This consistent finding sends a strong message to data policy-makers and practitioners about public dissatisfaction with existing data uses.
- 2. Concern about some data uses is high when compared with other, everyday concerns (eg migration or local crime).** Data uses that concern people include: data being used in unfair ways; commercial companies profiting from personal data; organisations tracking when, where and how people log on; and automated technologies being used to try to change people's behaviours. But concern about one data use is low: personal data being used to manage Covid-19.
- 3. Context is a defining factor** when it comes to public attitudes to data uses. Data uses in some contexts are more concerning than others. There are also differences in attitudes depending on the type of data use. **It is therefore not helpful to talk about public attitudes to data uses out of context.**
- 4. People want to know more about data uses, but the people who know most about them are the most concerned.** People who are more knowledgeable about data uses are more likely to have negative attitudes towards them. People also want information about what happens to their data – in its absence, they often imagine the worst. Put together, these findings tell us that **change is needed.**

Living With Data is a research project which aims to understand people's perceptions of data uses and how these could be improved. This document is a summary of the main findings from a survey carried out in late 2020 of people's attitudes to data uses. We describe our full findings and methodology in a detailed report which can be found here:

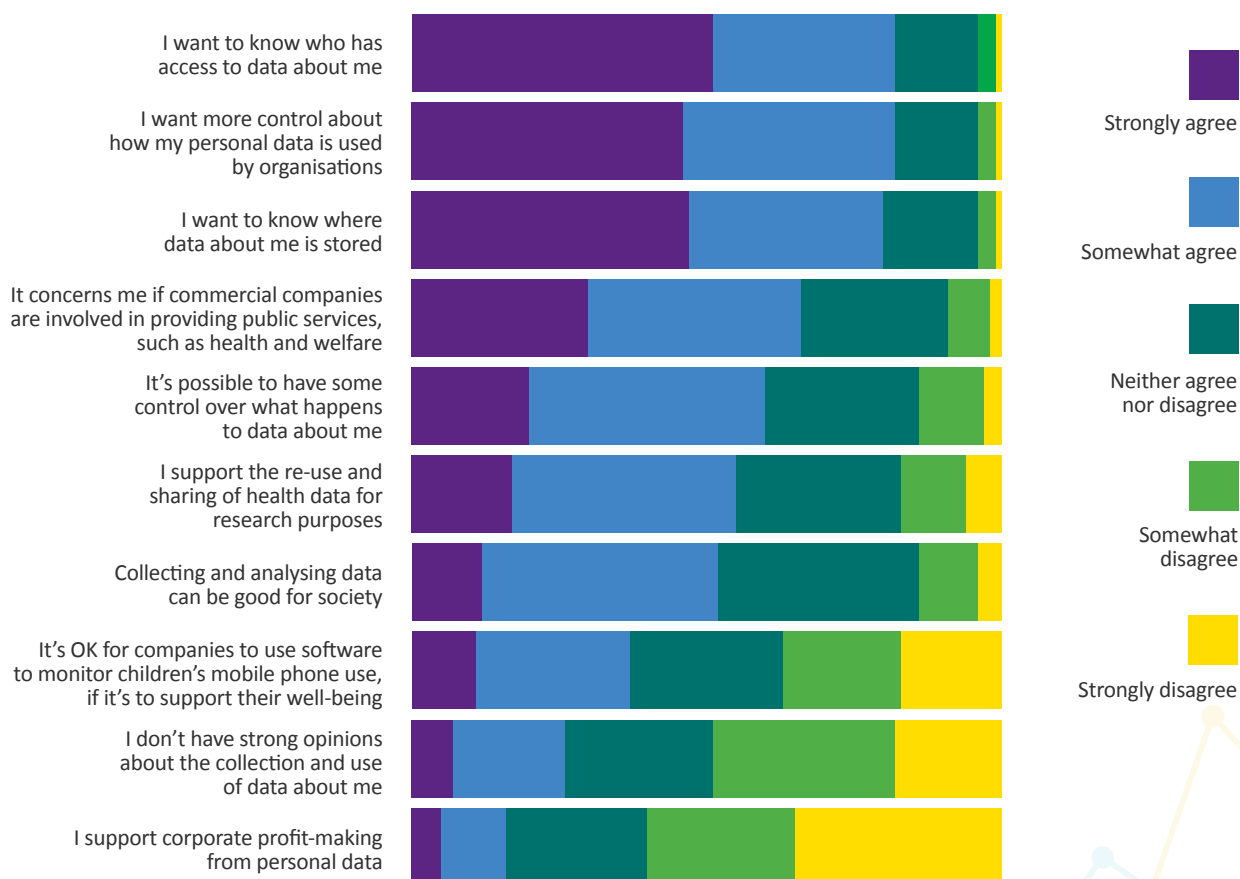
<https://livingwithdata.org/current-research/publications/>

MORE ABOUT FINDINGS

1. Public concern about certain data uses remains high

- Respondents want to know who has access to data about them (83% agree or strongly agree with a statement to this effect), they want more control over how their data is used by organisations (83%), and they want to know where data about them is stored (80%). They do not support corporate profit-making from personal data (60% disagree or strongly disagree). Only 26% of respondents agree or strongly agree with the statement 'I don't have strong opinions about the collection and use of data about me'. These findings suggest high levels of concern.
- 52% of respondents agreed or strongly agreed that collecting and analysing data can be good for society. This shows that whilst having some concerns about data uses, people can also recognise the benefits of data uses. However, only 12% strongly agreed – this low fraction is also indicative of concern.
- There were some differences in attitudes across groups, but differences were generally small.
 - Disabled people are less concerned about the re-use and sharing of health data for research purposes, but they are more concerned about commercial companies providing data-driven public services.
 - Women are more likely to want to know who has access to data about them than men, and to want more control over how their personal data is used by organisations.
 - People with no formal qualifications are more likely to indicate concern about data uses than people with high levels of qualifications. This was the biggest difference we found, with only around 40% of people with no formal qualifications agreeing with the statement 'Collecting and analysing data can be good for society', compared to around 60% of people with degrees.

Please indicate how strongly you agree or disagree with the following statements.



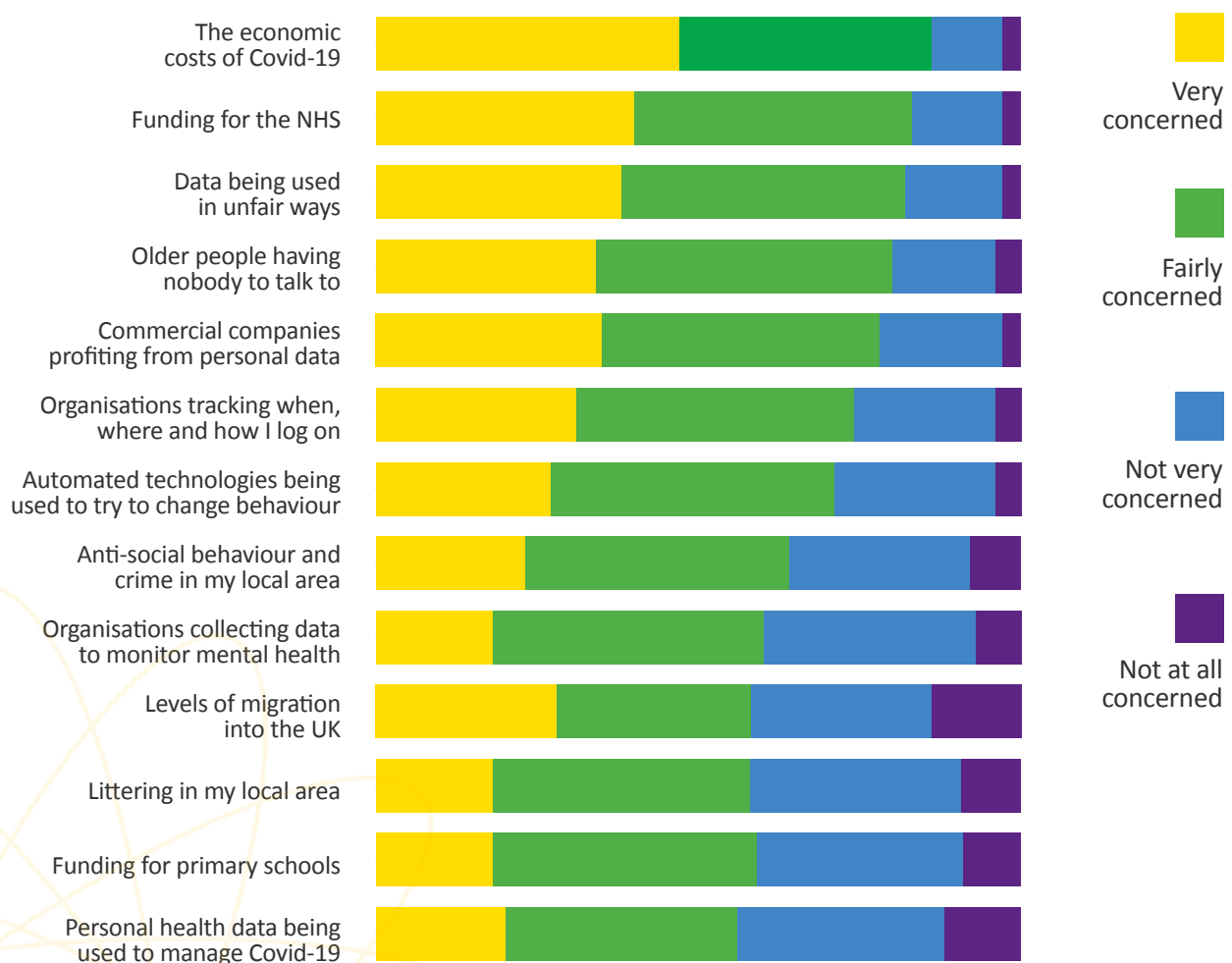
2. Concern about some data uses is high when compared with other, everyday concerns

We asked respondents how concerned they were about a range of issues (some of which were taken from Ipsos Mori's Most Important Issue tracker), to compare concern about data uses with other concerns they may have in their daily lives.

We found that:

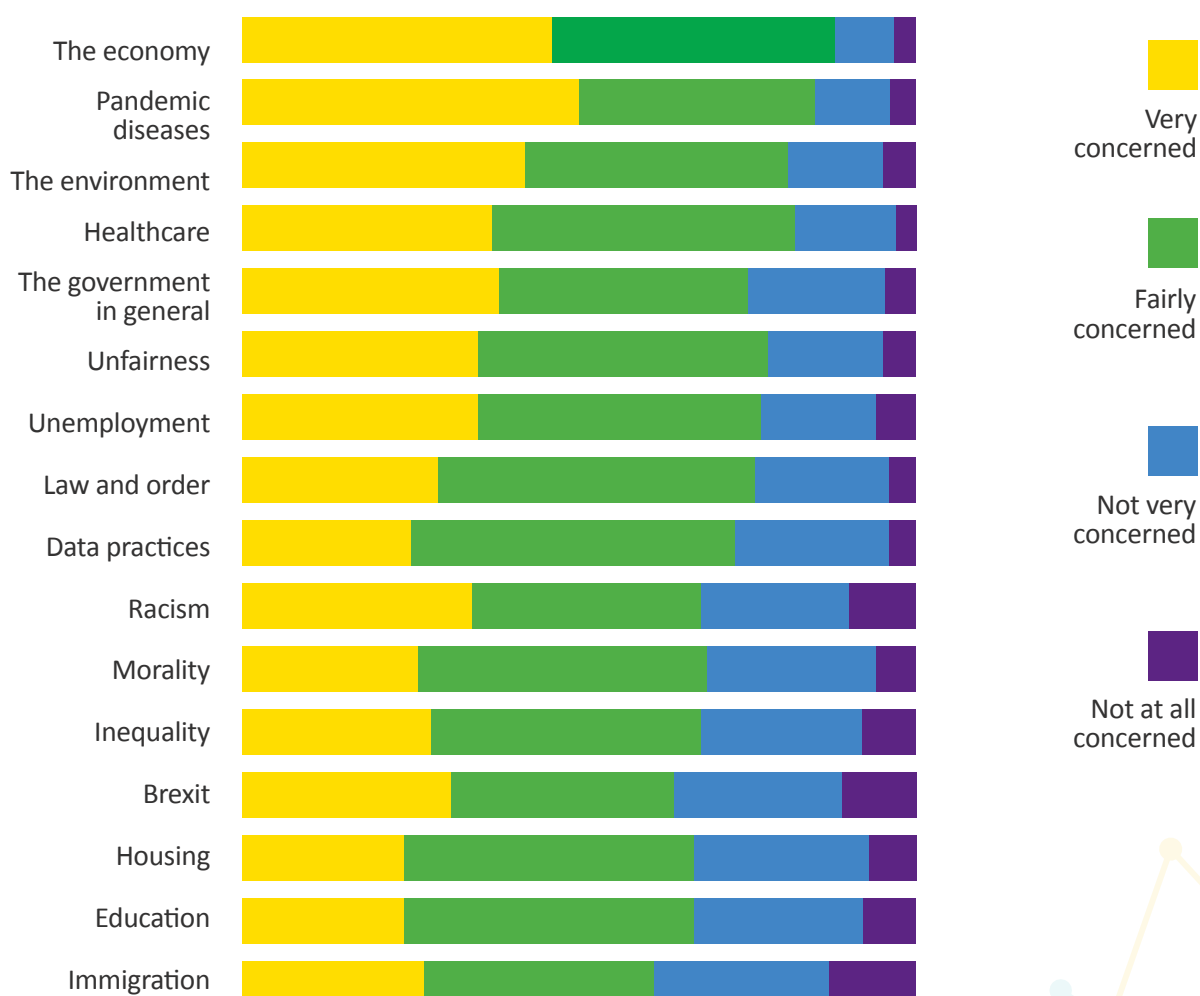
- In responses to questions about *narrow* concerns, data being used in unfair ways is a major concern. Only the economic costs of Covid-19 and funding for the NHS ranked more highly. Other specific data related concerns were also ranked highly. Commercial companies profiting from personal data; organisations tracking when, where and how people log on; and automated technologies being used to try to change people's behaviours were ranked the 5th, 6th and 7th biggest specific concerns respectively, in a list of 13.
- However, personal data being used to manage Covid-19 was the least concerning from the list of 13 issues shown in the figure below.
- In responses to questions about *broad* concerns, 'data practices' were ranked 8/15, less concerning than health and the environment, for example, but more concerning than potentially divisive issues like Brexit and racism. A large fraction of respondents described themselves as fairly concerned about data practices, but significantly fewer described themselves as very concerned.

In general in your daily life, how concerned are you about each of the following?



- However, when we asked respondents to tell us which three items from the list they were most concerned about, data practices was selected by the fewest people.
- These combined findings suggest that people are moderately concerned about the rather opaque issue of 'data practices', but it is rarely their main concern. In contrast, asking people their views on narrower aspects of data use unveils differing attitudes to different data uses, some of which are of considerable concern (eg data being used in unfair ways), some of very little concern (eg personal health data being used to manage Covid-19).
- Younger people, people of colour, LGBTQ+ people, and people in receipt of Universal Credit were more concerned about narrow data uses than other groups.
- These questions came after respondents had completed a survey about their attitudes to data uses, so it may be that they expressed more concern than they would in a different context.

In general in your daily life, how concerned are you about each of the following?



3. Context is a defining factor when it comes to public attitudes to data uses

From the above we can see that attitudes to data uses differ depending on the context. The context of data uses, therefore, influences people's attitudes towards them. We found that this was the case elsewhere in the survey, when we asked respondents how much they trusted different sectors or organisations to a) keep their data safe, b) gather and analyse data in responsible ways, and c) be open and transparent about what they do with data.

We found that:

- Respondents said they trusted public health care professionals – specifically their GPs and the NHS – most with their data, and they trust media and social media companies the least.
- Levels of trust expressed by respondents were consistent across the three data uses that we asked about (keeping data safe, gathering and analysing data in responsible ways, and being open and transparent about what is done with data).
 - 72% of respondents trust their GP a great deal or a moderate amount to keep their data safe, 71% to gather and analyse data in responsible ways, 69% to be open and transparent about what they do with people's data.
 - There were similar results for the NHS: 69%, 69% and 67% respectively.
 - In contrast, only 5% of respondents trusted social media companies a great deal with regard to all three of these activities.
 - Figures were also low for the BBC: 36%, 35% and 33% respectively.
- This consistency in degrees of trust across the three data uses suggests that trust in sectors and institutions strongly influences trust in the same sectors' and institutions' data uses. In other words, context is a defining factor when it comes to public attitudes to data uses. It is therefore not helpful to talk about public attitudes to data uses out of context.

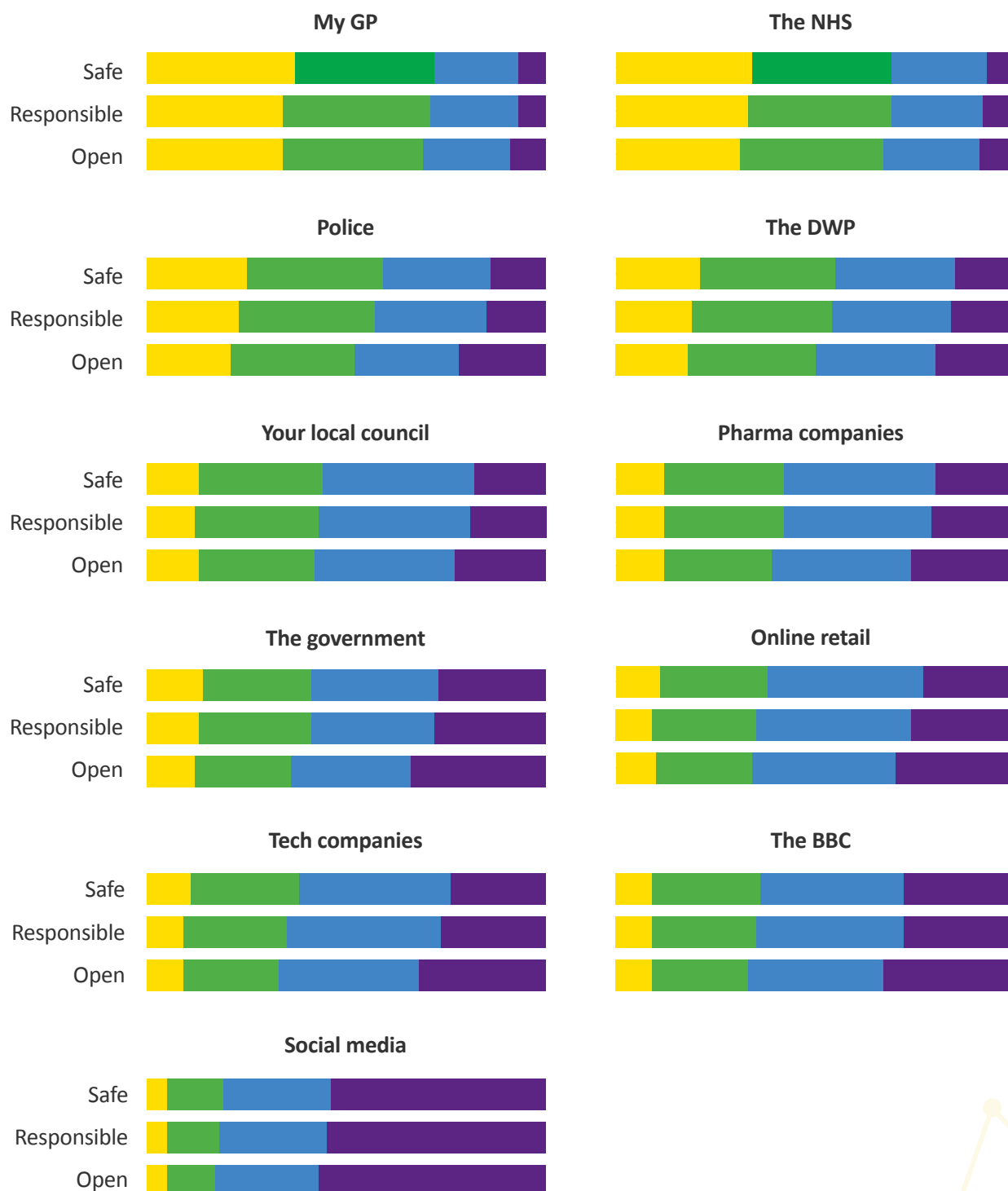
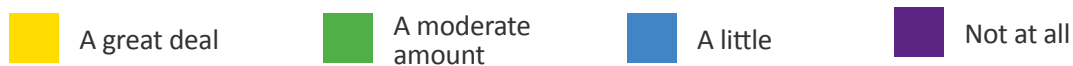
We asked respondents their views on different aspects of *specific public sector data uses*. The specific public sector data uses were:

- BBC experiments with individual control over uses of their personal data, **BBC Box**, a prototype device which pulls together data about what users watch or listen to and gives them control over who has access to this data, and **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;
- The **NHS Covid data store**, a national data store to hold data in one place to help national organisations responsible for coordinating the Covid-19 response;
- **Confirm Your Identity**, a DWP identity verification process for Universal Credit payments, which makes it possible to confirm identity online.

We found that:

- More than 3/4 of respondents were comfortable about their NHS patient data being added to the **NHS Covid data store**. Comments in free text fields confirmed that respondents supported the aims of the data store, but they also expressed more concerns about data sharing than respondents commenting on the other specific public sector data uses. More than half of comments expressed concern about the involvement of commercial companies in the NHS Covid data store. So respondents saw the value of using data to tackle Covid-19, but they were also concerned about potential future uses of the data in the data store.
- Concerns were often expressed as imagined future scenarios, most of which were negative, and many of which involved commercial organisations profiting from, leaking, misusing or selling data in the future. Out of all of the specific public sector data uses, negative imagined future scenarios like these were most common for the NHS Covid data store.

How much do you trust <organisation> to: keep data about you safe?; gather and analyse data about you in responsible ways?; be open and transparent about what they do with data about you?



- Respondents were largely comfortable with **DWP Confirm Your Identity**. Many respondents accepted the purpose of this data practice, and in free text fields, around 1/4 respondents indicated that they would use Confirm Your Identity because it is convenient.
- Some aspects of DWP Confirm Your Identity concerned respondents, such as those that might reinforce or lead to inequalities. They were not comfortable with the possibility that some groups of people might not be able to confirm their identities online because of their unusual residence histories or otherwise complicated lives, or that such people might have their access to Universal Credit delayed by the introduction of new forms of online identity verification.
- At the time of administering the survey, DWP was considering ways to make access to verified online accounts more secure, like extra checks designed to enhance security, which are in line with Industry Standards and used across a range of sectors. Such extra checks could include checking whether the same device is consistently used to log into accounts, whether people log in at similar times each day, whether the rhythm that the password is typed is consistent and whether the phone is swiped in a consistent pattern. We asked respondents how they felt about such checks. Respondents were least comfortable with DWP checking whether people logging into accounts swipe their phones in the same pattern as they normally swipe it.
- Most respondents were comfortable with **BBC Own It**. However, there were some aspects of the app that concerned them. Respondents were concerned about anonymous data about Own It users being collected by the BBC, but comfortable with the fact that this feature could be turned off.
- When asked about possible future models for sharing Own It data with university researchers in order to better understand children's mental health, the most popular model was no data sharing (43% of respondents chose this option) and the least popular model was sharing anonymised, individual level data (8%). Other options included sharing anonymised data (27%) and sharing anonymised, group level data (22%).
- Just 4% of people said they definitely would use **BBC Box** and around a quarter stated that they definitely wouldn't use it. The free text fields suggest that this may be down to a lack of interest, perceived need or demand for it.
- Lack of trust in the organisation undertaking the data practice was mentioned more frequently in relation to BBC Box than in relation to the other specific public sector data uses. Of the people who said they trusted the BBC a great deal to be open with what they do with users' data, 21% said they would definitely use a physical BBC Box, compared with just 3% of those who said they did not trust the BBC at all on this same issue.
- There were some differences in groups' attitudes to these specific data uses. For example:
 - Young people were much more positive about **BBC Box**, perhaps because they were better able to understand it.
 - People in receipt of Universal Credit were more receptive to **DWP Confirm Your Identity** automated checks than other respondents. This may be because DWP generates trust through transparency about its data uses, because they are familiar with such online systems, or it may be a result of 'digital resignation' (Draper and Turow 2019) – that is, feeling that they have no choice but to share personal data in order to access services. Our qualitative research is exploring this further.
- Although context is a defining factor, type of data use is also important. For example:
 - Respondents were broadly comfortable with **BBC Own It**, but they were concerned about anonymous data being gathered by the BBC and did not support possible future data sharing with university researchers, even if this might lead to better understanding of children's mental health.
 - Likewise, they were broadly comfortable with **DWP Confirm Your Identity**, but were concerned about the possibility of it disadvantaging certain groups, and about some of the possible extra checks that DWP was exploring. In free text fields, some participants note that some data uses feel 'creepy', even if they appreciate the reasons behind them.

4. The more people know about data uses, the more they are concerned about many of them

Are people who are more knowledgeable more or less concerned about data uses? Given widespread claims about the importance of more transparency about data uses (for example, by the Office for National Statistics, and the Information Commissioner's Office in the UK), the answer to this question matters.

To explore the relationship between (i) people's awareness of data practices and (ii) their attitudes towards them, we produced classifications of awareness and attitudes. We used latent class analysis to do this, based on our analysis of responses, rather than how respondents described themselves.

A detailed description of how we did this can be found in our full report:
<https://livingwithdata.org/current-research/publications/>

We found four clusters of awareness, which we call:

- **Knowledgeables** (25%): respondents who generally correctly identify true and false statements;
- **Believers** (38%): generally respond that the statements they are presented with are true;
- **Disbelievers** (14%): generally respond that statements are false, even when they are in fact true;
- **Don't knows** (23%): generally state that they don't know the answer to awareness questions.

We found four clusters of attitudes, which we call:

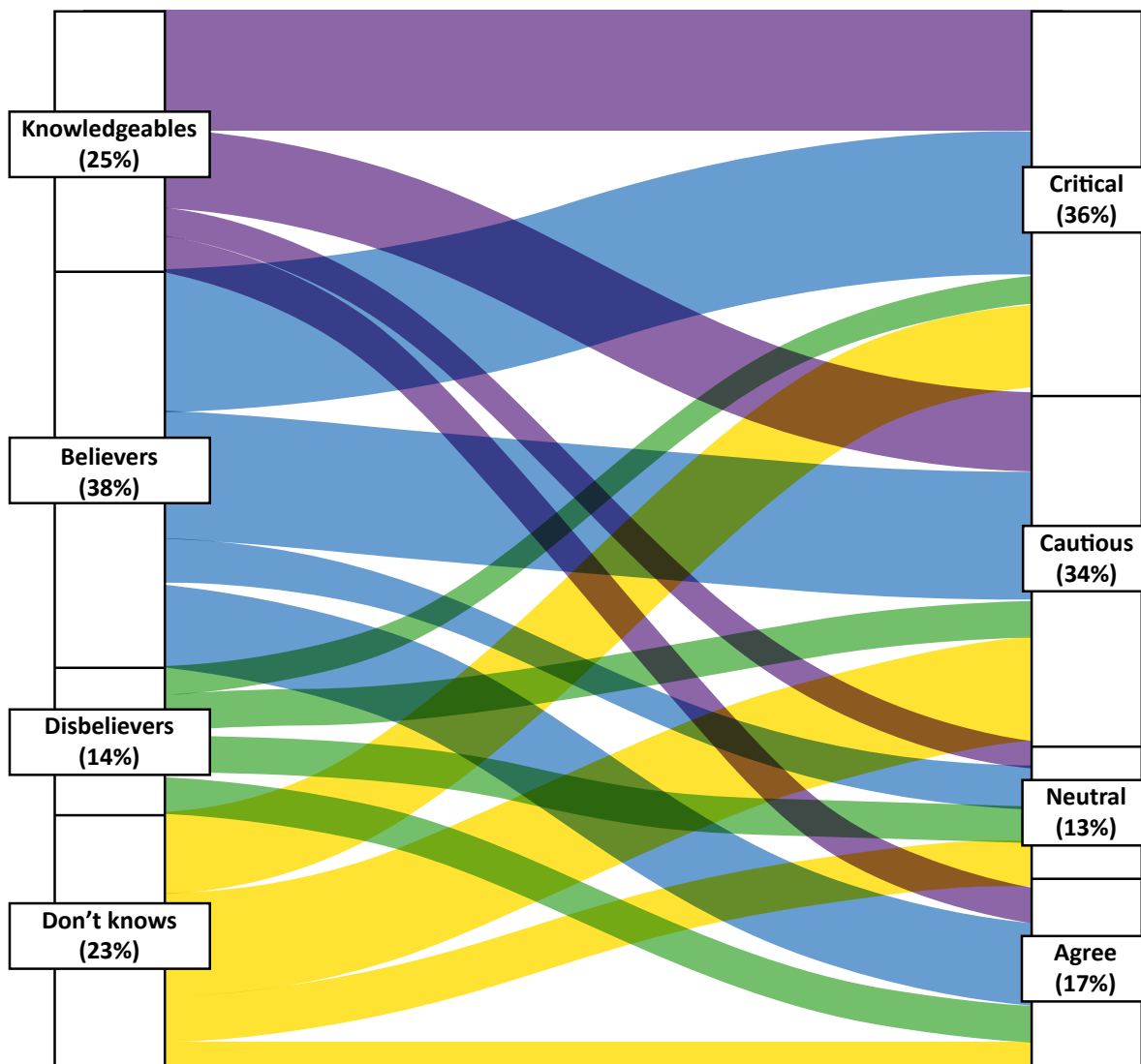
- **Critical** (36%): respondents who strongly disagree with some statements that are positive about data practices, and strongly agree with others that are negative about data practices;
- **Cautious** (34%): tend to agree or disagree in the same directions as the Critical group, but not strongly;
- **Neutral** (13%): overwhelmingly respond with 'Neither agree nor disagree';
- **Agree** (17%): overwhelmingly either agree or strongly agree with all statements.

Examining the relationships between the clusters, we found:

- **Knowledgeables** are more likely to be in the **Critical** group, and moderately more likely to be in the **Cautious** group. In other words, people who are more knowledgeable about data uses are more likely to have negative attitudes towards them.
- There is also a clear relationship between respondents who are **Neutral** in attitude (who generally neither agree nor disagree) and **Don't knows** in awareness (who generally stated that they didn't know the answer to questions). So people who are less confident about their knowledge of data uses are also less likely to hold strong opinions on them.
- People who are not aware of data uses (**Disbelievers**) were the most likely to be comfortable with the data uses presented to them, while more knowledgeable people were less comfortable and more critical people were much less comfortable.
- **Knowledgeables** and **Critical** respondents are less likely to use or less comfortable with the specific public sector data uses presented to them.
- **Knowledgeables** were least likely to report that they were surprised by the data uses described, but they were no more likely than other groups to report that they understand them.
- There are some differences in membership of the different clusters according to respondents' demographic characteristics, but these are generally small. Younger people are much more likely than older people to be in the **Agree** attitude cluster – this is the largest difference of all in relation to the attitude classification.

In summary, the more people know about data uses, the more they are concerned about many of them. When we put this together with evidence which shows that people want to know about what happens to their data, this tells us that change is needed, because when people know about data uses, they don't like what they know.

The illustration shows the relationship between knowledge clusters (on the left) and attitude clusters (on the right). The bars connecting knowledge clusters to attitude clusters show the overlap between them.



CONCLUSIONS

Public concern about certain data uses remains high.

Survey respondents indicate high levels of concern with particular aspects of data use and data uses in particular contexts. For example, across different parts of the survey, we found that people do not like commercial companies profiting from personal data, especially if this data was originally gathered for pro-social purposes such as public health. Even where we found statistical support for data uses like the NHS Covid data store, in written comments respondents expressed significant concern – in this case, about commercial companies profiting from or misusing data. Similar high levels of concern have also been identified in previous research. This consistent finding communicates a strong message to data policy-makers and practitioners about public dissatisfaction with existing data uses.

It is important to note that in our survey, we sought to explore public attitudes to data uses, whether positive or negative. We worded our questions carefully and made sure to include a balance of statements about the potential benefits as well as the potential harms of data uses. We did not look for concerns alone. Given this, it is noteworthy that we find high levels of concern about data uses across all demographic groups: it is not the case that only groups that may be disadvantaged or harmed by certain data uses are uncomfortable with them.

Context is a defining factor when it comes to public attitudes to data uses.

Some data uses are more concerning than others. Commercial companies profiting from personal data, organisations tracking when, where and how people log on, and automated technologies being used to try to change people's behaviours are more concerning than personal data being used to manage Covid-19. It is therefore not helpful to talk about public attitudes to data uses out of context, or to assume that public views of data uses will be consistent across contexts. Rather, we need to be attentive to contextual specificity. Furthermore, complexities arise when less concerning data uses come together with those that are more concerning. In the case of the NHS Covid data store, respondents support its pro-social, public health aims and are simultaneously concerned about the involvement of commercial companies and the lack of clarity about the ways in which these companies may have access to or profit from personal data in the future.

Within different contexts, there are differences in attitudes depending on the type of data use in question. Respondents were broadly comfortable with BBC Own It, but concerned about some details. Likewise, they were broadly comfortable with DWP Confirm Your Identity, but they had some concerns about some of the possible extra checks that DWP was exploring, such as checking whether people logging into accounts swipe their phones in the same pattern as they normally swipe it, even though these are designed to enhance security.

It is also important to note that context influences trust. Therefore, trust issues go beyond the specific data use. People trust the health care sector a lot more than they trust technology, media and social media companies, regardless of the type of data use. So high trust in GP and NHS data uses is in danger of being diminished by the involvement of less trusted parties, such as technology companies, in health data initiatives like the NHS Covid data store. Generally, degree of trust in data use results from attitudes to organisations or sectors more generally, so to address data trust issues, change needs to go beyond data uses and to focus on trustworthiness more generally.

People want to know more about data uses, but the more they know, the more they are concerned about many of them.

People who are more knowledgeable about data uses are more likely to have negative attitudes towards them. How we interpret this finding is important. We should not conclude that we therefore shouldn't tell people about data uses! People want information about what happens to their personal data: they want to know who has access to data about them and where data about them is stored. In the absence of clear information about data uses, people imagine negative scenarios in which their data may be sold, leaked or misused, and this leads to concern and distrust. Put together, these findings tell us that change is needed.

RECOMMENDATIONS

Future data policy-making and data practice should be informed by our findings, for example:

- **Consider NOT engaging in practices** that consistently concern the public, for example sharing data originally intended for health or other pro-social purposes in ways which enable commercial companies to profit.
- **Communicate clearly and fully** about existing data uses and possible future uses, especially where such uses go beyond what was originally intended, because in the absence of clear information, people imagine negative and concerning future data use scenarios. Doing this requires investment of time and effort.
- **Change practices.** The people who know most about data uses are the most concerned about them. If data uses continue unchanged, the public will continue to be concerned, regardless of clarity of communication.
- **Consult the public**, particularly those affected by data uses, about what changes to data uses are needed. Engage in a genuine dialogue with the public about data uses on a case-by-case basis, because public views of data uses differ based on context.
- Understand that **some trust issues go beyond the type of data use** and instead result from attitudes to organisations or sectors more generally. To address data trust issues, fundamental changes may be needed that go beyond data uses and practices.

ABOUT THE SURVEY

Living With Data is a research project funded by the Nuffield Foundation, which aims to understand people's perceptions of how data about them is collected, analysed, shared and used, and how these processes could be improved. We use the term 'data uses' as shorthand for these processes.

We carried out a survey in late 2020 of people's attitudes to data uses. We used the web survey platform Qualtrics to recruit and administer the survey to 2000 adults in the UK. The sample was nationally-representative, with additional

recruitment (or 'boosts') of people born outside the UK, LGBTQ+ people, and people in receipt of Universal Credit, to ensure their voices could be heard in our research.

The survey consisted of seven parts, in which we asked about: personal characteristics; internet usage; awareness of data uses; general attitudes to data uses; trust in institutions' data uses; attitudes to specific public sector data uses in the BBC, DWP and NHS; and concerns about data uses as they compare with other concerns.

livingwithdata.org



Living With Data is funded by the Nuffield Foundation, but the views expressed are those of the authors and not necessarily the Foundation.

