



UK public unhappy with the way their personal data is managed

Findings from a survey of public views on data management models

Helen Kennedy, Todd Hartman and Robin Steedman, University of Sheffield
Rhianne Jones, BBC Research and Development

SUMMARY

- A consistent finding is that the UK public dislikes the current, status quo model for managing data, in which commercial organisations control personal data in return for the digital services they provide.
- The public prefers approaches that give individuals control over their personal data, that include oversight from regulatory bodies and/or that enable opting out from data gathering.
- Variations of data trusts were preferable to the status quo, but not as widely preferred as models involving personal control, regulatory oversight or the ability to opt out.
- The public prefers all credible alternatives to the status quo.

WHY ARE DATA MANAGEMENT MODELS IMPORTANT?

The widespread collection and use of digital data is said to have wide-ranging effects: benefits such as more effective service provision, and harms such as more surveillance, less privacy, and new forms of inequality and injustice. Research has shown that there is growing concern about the possible negative consequences of the use of data-driven services and platforms. Low levels of public trust have been identified (Open Data Institute 2019a), something which has been described as a ‘data trust deficit’ (Royal Statistical Society 2014). Awareness of this data trust deficit, combined with high profile failures to protect people’s personal data from exploitation or misuse, has led to a growing consensus on the need for responsible data practices.

In the UK, the government has identified public trust in data as a critical enabling factor for the success of the UK’s National Data Strategy (Department for Culture, Media and Sport 2018). The EU General Data Protection Regulation (GDPR) of 2018 provides a legal motivation to improve data practices. Under

GDPR, individuals have new rights with regard to their personal data, including right of access and portability. These factors have led to experimentation in alternative approaches to the management of personal data. Although most online companies continue to collect and have significant control over data about their users, other approaches to data collection and storage are being considered, which could be better for individuals and society. These include Personal Data Stores, data trusts and other forms of responsible data stewardship.

It is vital that public views are factored into decisions that will shape the future of the data economy, including in relation to data management models. Interest in how the public perceives data practices has begun to grow, and a number of recent surveys and polls have looked at public attitudes towards these things. However, there has been no research on what the public think about alternative approaches to data management. The survey we carried out and discuss in this report fills that gap.

OUR SURVEY OF ALTERNATIVE DATA MANAGEMENT MODELS

2,169 respondents living within the UK completed our online survey from an opt-in Qualtrics panel in May 2019. We collected data from respondents from across all regions of the UK, including Northern Ireland. Respondents were diverse in relation to gender, age, ethnicity, educational attainment, employment status, and household income. In the survey, we examined what respondents thought about

eight models for managing personal data. Each model was based upon approaches to data management that were being considered in various forms at the time of administering the survey, including Personal Data Stores, data trusts and data co-operatives. The models and descriptions used in the survey are summarised in Table 1.



Table 1: List of data management models evaluated in the survey

Name	Description
Personal Data Store	You are given a secure place to collect, store and manage the data about you which has been collected by other services. This is called a Personal Data Store, or PDS . You have access to this data, and you can decide who else can access this data, how they can use it and under what circumstances. The purpose of the PDS is to give you personal control over your data, which you can manage in a secure way.
Responsible Independent Party	You are given a way to nominate a responsible independent party to oversee collection, storage and access of your personal data. They have legal responsibilities to look after your data. In line with your wishes, the nominated party can make decisions on your behalf about who accesses your data, what they can do with it and under what circumstances. You have a say over what happens to your data, but you are not personally responsible for looking after it.
Responsible Independent Organisation	Responsible independent organisations manage your data in different contexts (eg one for health data, one for finance data, etc). These organisations make decisions about who can access your data, what they can do with it and under what circumstances. They have legal responsibilities to manage access to your data in ways that represent the interests of all parties involved.
Status Quo (described as a Digital Service)	You sign up to a new digital service (eg an online shop) that collects and uses your data. You are asked to agree to terms of use and a privacy policy beforehand. These describe how the service will collect, store and manage data about you. You are given settings you can alter, but you are not able to change or negotiate these terms or see how your data is used. This approach gives services control over your data (this is what usually happens now).
Data Co-operative	You become a member of a data co-operative that manages the collection and storage of its members' data and is accountable to its members. As a member, you can put yourself forward to sit on a board of representatives and make decisions about who has access to members' data, how it is used and under what circumstances. Or you can vote for other co-operative members to do these things. The purpose of the data co-operative is that your data is managed collectively, by the people whose data is in the co-operative.
Public Data Commons	You access data online about your area and community using an open data platform that is accessible to all citizens under commons law. This is called a public data commons . The data commons collects, stores and manages access to open data which can be used for various purposes. Everyone can access and use this data, in line with the commons' rules of engagement. The purpose of the public data commons is to make data accessible so everyone can benefit from it.
Regulatory Public Body	You have been given the details of a new regulatory public body that oversees how organisations access and use data, acting on behalf of UK citizens. This public body provides oversight over how organisations collect, store and use personal data. It can hold organisations accountable for misuse (eg fine organisations when they breach terms of use). The purpose of the regulatory body is to ensure that personal data are collected, stored and used in legal and fair ways.
Opting Out (described as a Data ID Card)	You have the ability to choose whether to opt out of online data collection, storage and use – this is called managing your data preferences. Your data preferences are stored on a data ID card . You can use this card to log onto online sites. The card automatically opts you out of data collection, storage and use according to your preferences and whenever this is possible. The purpose of the data ID card is to give people the option of opting out of having their data collected.

We examined views on the eight data management models using three different methods:

1. Respondents were asked to rate randomly selected data management models (presented one at a time) using a scale ranging from 0 ('poor') to 10 ('excellent').
2. Because assigning a numeric value on an 11-point scale can be difficult, we also assessed preferences by presenting respondents with randomly generated pairs of models and asking them to choose the preferred option.
3. We randomly combined multiple features (type of data, who controls data and how, uses and beneficiaries of data) into scenarios and asked respondents to express preferences for scenarios generated in this way, in order to assess the relative effect of each feature on preferences for data management models.

FINDINGS: THE PUBLIC PREFERS ALL ALTERNATIVES TO THE STATUS QUO

What was most striking about the results is that respondents preferred **all other models to the status quo**, described in the survey as a ‘digital services model’ that ‘gives services control over what happens to your data’, the prevailing online model at the time

of writing. With an average rating of 4.9 out of 10, this suggests that respondents are not happy with the current ways in which services and organisations control data.

How the public rates models for managing their personal data

They dislike the status quo

Give services control over my data
★ ★ ★ ★ ★ 4.9/10

Their top three preferred alternatives

Control over my personal data
★ ★ ★ ★ ★ 7.7/10

Oversight from regulatory bodies
★ ★ ★ ★ ★ 7.6/10

The option to opt-out from data gathering
★ ★ ★ ★ ★ 7.5/10



Of the eight data management models that we presented to respondents, the most preferred approach was the Personal Data Store (PDS), described in the survey as ‘a secure place to collect, store and manage the data about you which has been collected by other services’. The Personal Data Store would give individuals control over their personal data, and responses to other questions in the survey about views on data uses suggest that this may

be why this model was highly rated. For example, 86.9% of respondents agreed with the statement ‘I want more control over how my personal data is used by organisations’, and 89.0% agreed with the statement ‘I want more control over my personal data.’ This finding is in line with previous research which has also identified the importance of control (eg Digital Catapult 2015).

Table 2. Mean ratings on a scale from 0 to 10 for each data management model

Model	Mean rating
Personal Data Store	7.7
Regulatory Public Body	7.6
Data ID Card (with clear opt-out options)	7.5
Responsible Independent Organisation	6.4
Public Data Commons	6.3
Responsible Independent Party	6.2
Data Co-operative	5.9
Digital Service (Status Quo)	4.9

The next most highly rated model after the PDS was one which involves oversight by a regulatory public body which would oversee ‘how organisations access and use data, acting on behalf of UK citizens’ in order to ‘ensure that personal data are collected, stored and used in legal and fair ways’. Elsewhere in the survey, we asked respondents who they would like to see provide new data-driven services ‘for the public good’ and most selected governmental or publically-funded organisations. Thus oversight of data-management by a public regulatory body is a strong preference.

The high rating of this model by respondents suggests a preference for legally enforceable safeguards alongside the personal control of data offered by the Personal Data Store. A Royal Statistical Society survey found that ‘there is more support for the government preventing misuse of personal data than there is an appetite to have personal control over this’ (2014, p.3). In contrast, we found a desire for both governance and personal control, which suggests that both are important principles in data management for the UK public. Both approaches would result in uses of data that are preferable to the status quo.

The third most highly rated model would allow people to opt out of having their data collected. We described this as a ‘Data ID Card’, to give material form to a model for opting out of data collection. The relatively high rating of this model also points to the importance of individual control over data for our respondents, as well as indicating a strong dislike of the status quo.

Data trusts, co-operatives and commons-based data management models all involve trusted parties overseeing, managing and stewarding data on behalf of individuals and communities. In our survey, we explored these models as follows:

- the data co-operative, which manages the collection and storage of its members’ data, is accountable to its members and is governed by a board of representatives constituted by its members;
- the data commons, similarly collectively motivated, which enables online access to community data which can be used for various purposes and for the benefit of all; and data trusts;
- two types of trust:
 - a trust governed by a responsible independent party which makes decisions on behalf of data subjects about who accesses data, what they can do with it and under what circumstances,
 - a trust governed by multiple responsible independent organisations which manage different types of data in different contexts (for example, one for health data, one for finance data, and so on) and represent the interests of all parties involved.

These four ‘trust-like’ models were preferable to the status quo, but not as widely preferred as models involving personal control, regulatory oversight or the ability to opt out. They all had lower mean scores than those that offer personal control or regulatory oversight.

The second method we used in the survey presented respondents with a randomly generated pair of models and asked them to choose the preferred option. The top three preferred models using this method were the same as those reported above: the Personal Data Store, opting out and oversight by a regulatory public body, in that order of preference. There was a 30 percentage point increase in selecting the top three data management models compared to the current, status quo approach. In other words, respondents were 30 percentage points more likely to choose the top three models. This is a significant number, both statistically and substantively. As with our first ratings method, trust-like models were preferable to the status quo, but less preferable than those based on personal choice and control and regulation.

The third method we used in the survey randomly combined multiple features into scenarios and asked respondents to express a preference for one of two scenarios generated in this way – an example is given below. We did this to compare the significance of a number of features of data handling scenarios, such as types of data, control and rights, uses of data and related benefits. The single most important feature influencing responses was the locus of control over the data. When the data is controlled by ‘you’ instead of a commercial organisation, we saw the largest increase in the probability of selecting that scenario. In other words, controlling their own data was really important to respondents. Other significant attributes relate to uses and beneficiaries of the data. Respondents were more likely to prefer scenarios in which data would be used for personal insights or to benefit society than for profit. Respondents preferred scenarios that provided them with the right to access their personal data, have more control over it, and know that it is secure. But compared to other features discussed above, individual rights contributed only small effects to the overall decision.

Table 3: Example scenarios from our third method

	Option A	Option B
In this scenario the data is	Medical data	Financial data
The data is controlled by	You	A trustee like a city council or the government
You will be able to	Have full control over what happens to it	Know what data is held about you, by whom and what they do with it
The data will be used for these reasons, and generate these benefits	So you can get insights and value from your personal data	So an organisation can use your data to benefit the public

We found that existing knowledge about relevant issues was a significant predictor of preferences in relation to four models. More knowledgeable respondents preferred approaches that offered more control and/or oversight over personal data by a regulatory public body than less knowledgeable respondents, who rated the status quo higher. While this effect was significant, it was relatively small (about

a half point difference on a 10-point scale). In other words, this mattered, but not a great deal. Age also had a significant impact on evaluations of the status quo. Younger respondents rated the status quo higher than those who were older than 34. Apart from these two findings, there were no other clear differences in data management model evaluations by demographic subgroups within the sample.

CONCLUSIONS AND RECOMMENDATIONS

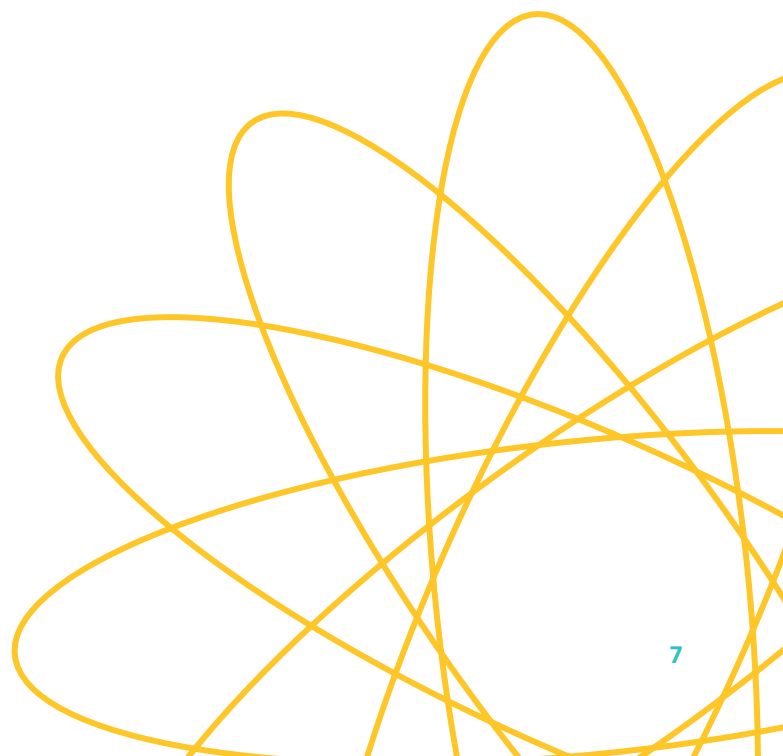
Our findings suggest that new approaches to data management are urgently needed, because there is a strong desire from the public for an alternative to the status quo. A consistent finding is that people dislike the status quo, in which commercial organisations control personal data in return for the digital services they provide. Respondents preferred approaches that give individuals control over their personal data, that include oversight from regulatory bodies or that enable opting out from data gathering. All credible alternatives—including a public data commons, a data co-operative, oversight by a responsible independent organisation or party—were preferable to the status quo.

These findings were consistent across different methods used in the survey: asking respondents to rank models on a scale, to choose a preferred model from a randomly generated pair; and to choose a preferred scenario from a randomly generated pair made up of different features. Policy-makers and organisations that handle personal data need to accept that the status quo is not sustainable, if they wish to factor public views into new developments and decisions that will shape the future of the data economy. New approaches need to give individuals control over their personal data and include oversight from regulatory bodies.

However, not all alternative approaches to data management were rated equally. Respondents expressed a greater preference for some than for others. Data trust-like models—a public data commons, a data co-operative, oversight by a responsible independent party or organisations—were ranked below PDS, regulatory and opt out models. These findings were consistent across different methods used in the survey. We cannot therefore conclude that there is a ‘huge appetite’ for data trusts amongst the public, as the ODI concluded from their data trust pilots with organisational stakeholders (ODI 2019b). Further research is needed to explore the reasons for this finding.

The implementation of alternative data management models will require investment of resources, to support technical development, testing and iteration, and public consultation. In addition, we need to understand why people prefer particular data management models and the extent of public understanding of differences across models. To further advance understanding of public views about data management models, more research is needed. This also requires investment in resource. In short, to ensure public support for data-driven developments, it is clear that:

1. **New approaches to data management are urgently needed**, which give individuals control over their personal data and include oversight from regulatory bodies.
2. **Investment of resources is required**, to support technical development, testing and iteration, and public consultation, if the implementation of alternative data management models is to be successful.
3. **More research is needed** to further advance understanding of public views about data management models.



REFERENCES

Department for Digital, Culture, Media & Sport (2018) Data mobility: The personal data portability growth opportunity for the UK economy. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/755219/Data_Mobility_report.pdf.

Digital Catapult (2015) Trust in personal data: A UK review. Report by Digital Catapult. Available by request from Digital Catapult.

Open Data Institute (ODI) (2019a) Attitudes towards data ethics. Available at: https://theodi.org/article/nearly-9-in-10-people-think-its-important-that-organisations-use-personal-data-ethically/?mc_cid=8fab3ef311&mc_eid=0e51d41637.

Open Data Institute (ODI) (2019b) Huge appetite for data trusts. Available at: <https://theodi.org/article/huge-appetite-for-data-trusts-according-to-new-odi-research/>, 15th April 2019.

Royal Statistical Society (2014) Trust in data and attitudes toward data use/data sharing. Available at: <https://www.statslife.org.uk/images/pdf/rss-data-trust-data-sharing-attitudes-research-note.pdf>.

ABOUT THE RESEARCH AND THE REPORT AUTHORS

The research formed part of an ongoing collaboration between the University of Sheffield's Digital Society Network (<https://www.sheffield.ac.uk/faculty/social-sciences/digital-society-network/home>) and BBC R&D (<https://www.bbc.co.uk/rd>). It was made possible by funding from the Arts and Humanities Creative Economy Engagement Fellowship Scheme (<https://ahrc.ukri.org/innovation/knowledgeexchange/cultural-engagement-fund/>), BBC R & D and the University of Sheffield.

A longer version of the research report can be found here: <https://livingwithdata.org/previous-research/views-on-data-management/>.

Helen Kennedy is professor of digital society at the University of Sheffield, director of the Digital Society Network and of the Living With Data programme of research hosted there (<https://livingwithdata.org/>).

Todd Hartman is senior lecturer in quantitative social science and a political psychologist, with international expertise in survey research and experimental design, also at the University of Sheffield.

Robin Steedman is a postdoctoral researcher interested in diversity and inequality in the media, formerly of the University of Sheffield and now of Copenhagen Business School.

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