Department for Business & Trade

Scenarios Report: The Future of Smart Data (2028)

July 2023

A joint project between the Centre for Data Ethics and Innovation (CDEI) and Department for Business, Energy, and Industrial Strategy (BEIS)

The Centre for Data Ethics and Innovation (CDEI) worked with the Smart Data team in the Department for Business, Energy, and Industrial Strategy (BEIS) to identify the features of ethical and trustworthy Smart Data schemes. The aim of this paper is to pull together the results of the CDEI's Scenario Planning Workshop, which was run in October 2021, outlining potential future evolutions of Smart Data schemes to inform future thinking.

The intended audience of this paper is: government departments, regulatory bodies, data holders, Authorised Third Parties, and consumer interest groups, particularly those that have been involved in the Smart Data Working Group (SDWG).

This is a research paper, and not intended to be a statement of government policy in this area.

Table of Contents

Introduction

The Centre for Data Ethics and Innovation (CDEI) is working with the Department for Business, Energy and Industrial Strategy (BEIS) to develop guidance on how to develop Smart Data schemes that are ethical and trustworthy. This project was run in three phases:

- *Phase 1:* Semi-structured interviews to gain insights on the features of ethical and trustworthy Smart Data schemes.
- *Phase 2:* Scenario planning workshop to understand how Smart Data schemes may evolve in the coming years.
- *Phase 3:* Developing an implementation guide for actors across the Smart Data ecosystem, responsible for designing and implementing both Smart Data schemes and services.

This report is the result of phase 2 of the project: the CDEI-BEIS scenario planning workshop that was held in October, 2021. Scenario planning is a widely used futures method that seeks to understand plausible, possible and probable futures for a given phenomenon.[1] It does so by bringing a group of expert individuals together to discuss the different trajectories a phenomenon could take under different circumstances. Importantly, scenario planning *does not seek to predict the future*. Instead, it is a heuristic for exploring complex and long-term issues, which leads participants to question their previously held perspectives.

The specific focus of the CDEI-BEIS Smart Data scenario planning workshop was to consider the different ways in which Smart Data could develop in the years leading up 2028 (far enough in the future to allow for current legislative initiatives, but close enough to allow specific questions). This report outlines the findings of this exercise. The remainder of this report is structured as follows: section 1 outlines the scenario planning methodology and more detail on the rationale for choosing this; section 2 outlines the three scenarios that were created as part of this exercise; finally, section 3 reflects on the findings that emerged from the scenarios and subsequent discussion. This includes an analysis of the desirable and undesirable elements within each scenario, as well as the implications for the future development of Smart Data schemes.

[1] 'Futures', or Futures studies, refers to different approaches to thinking about the future and exploring factors that could give rise to possible and probable future characteristics, events and behaviours. https://www.gov.uk/government/groups/futures-and-foresight

Section 1. Methodology

The CDEI-BEIS Smart Data scenarios workshop sought to consider the different ways in which Smart Data could develop in the years leading up to 2028. We hosted 20 key stakeholders from across the Smart Data ecosystem, with representatives from the public, private and third sectors. Attendees included regulators, data holders, Authorised Third Parties[2], and interest groups representing consumer rights (particularly those of consumers with vulnerabilities).[3] The session was broken down into three sections: an introduction to the CDEI Smart Data project and futures methodologies; breakout sessions to develop three plausible scenarios for the future of Smart Data; and a group reflection on the governance consequences of the scenarios.

To form the starting point for scenarios in the breakout sessions, we focused on two 'axes of uncertainty' that would act as the cornerstones of the scenarios: public trust in data sharing and level of data portability (the ability for people to move data about them). The different combinations of these two axes (e.g. high public trust, high data portability; low public trust, high data portability) provide the starting point for creating four unique scenarios (see figure 1). These axes were used to create four potential futures as a starting point for discussion. Within three of these four scenarios, we asked a set of standardised questions to prompt stakeholders to discuss and develop different aspects of the scenario that they were creating (figure 2). These questions were developed from our phase 1 interviews and sought to assist in developing holistic scenarios that were relevant for the key questions considered by the CDEI in its project and the Smart Data team more broadly. Due to resourcing constraints, the workshop focused on three of these scenarios.

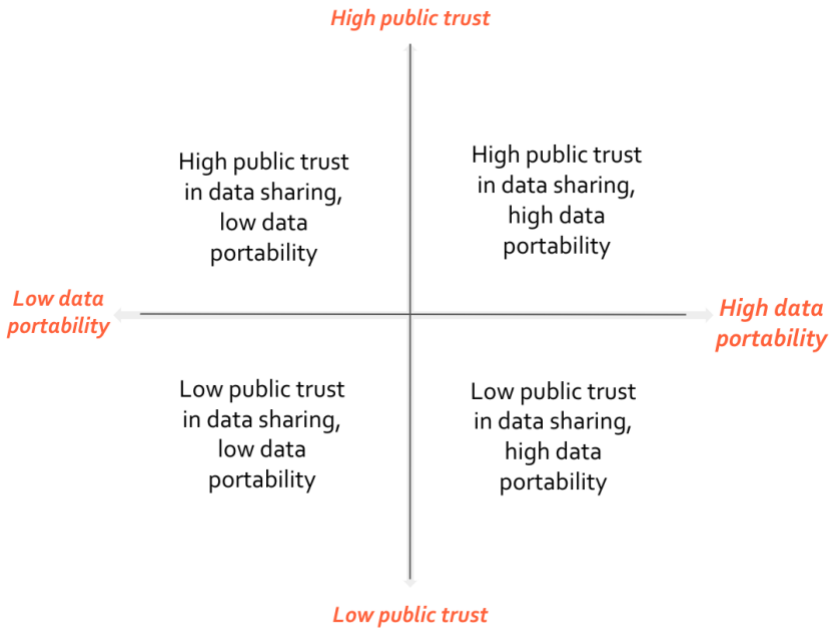

Figure 1: Axes of uncertainty

[2] Following consultation with stakeholders from across the Smart Data ecosystem, this paper uses the term "Authorised Third Parties" in place of TPPs, in order to be as clear as possible for consumers that these are authorised services. The term TPPs is still used in Open Banking.
[3] More details on the attendees of this workshop can be found below in Appendix 1 - Attendees.

The workshop broke into three breakout sessions to discuss three of the four scenarios developed above. This ensured that each of the three sessions was adequately represented by stakeholders from the public, private and third sectors. Participants were provided with prompt questions - outlined in Figure 2 - to guide discussions. The scenarios, as presented below, reflect the answers provided by participants.

One limitation of this approach was that each scenario was considered by only one group of stakeholders. Likewise, stakeholders were only able to explore one scenario. This could impact the consistency of approach, and the balance of opinions, between the different scenarios.

The scenario planning approach allowed participants to think expansively about the different plausible futures that could emerge for Smart Data and the features that these futures could be defined by. This is important for Smart Data, given the multitude of governance options available. Understanding which features of a given scenario were positive and negative, and what precipitated these features, allowed stakeholders to think about tangibly different governance options. This helps avoid either thinking too narrowly about governance options and not considering potential impacts on the one hand, and focusing on abstract ethical discussions on the other.

The year 2028 was chosen as it was far enough in the future to allow for current legislative initiatives by BEIS, and complementary secondary measures to have matured. At the same time, 2028 was near enough in the future to be able to focus on the specific questions we are seeking to address, without having to worry about longer term shifts that would impact the scenario.

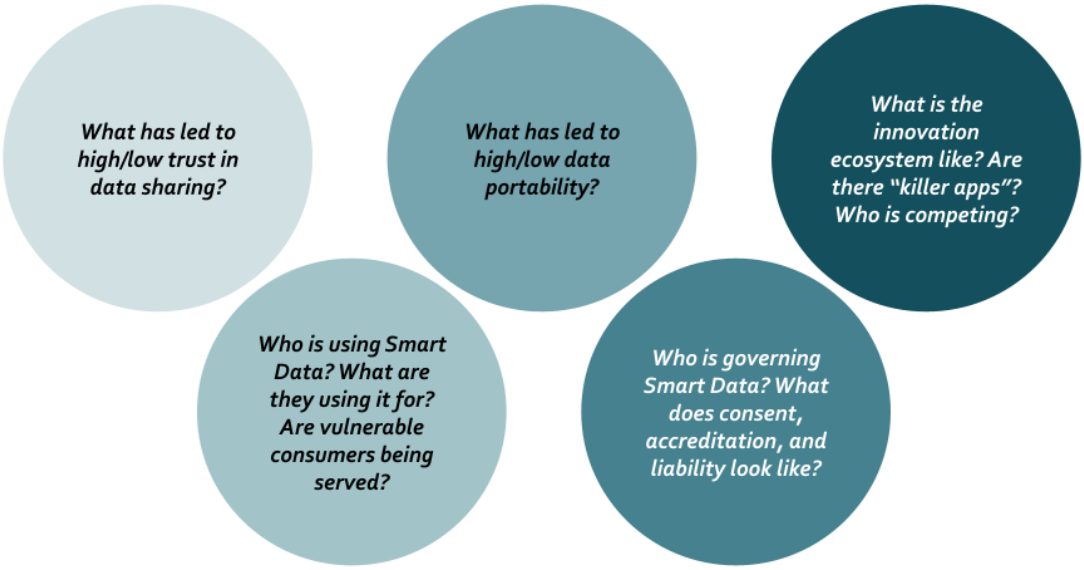

Figure 2: Discussion questions posed to participants in the breakout sessions

Section 2. Scenarios

These scenarios are not official HMG policy or recommendations

These scenarios are intended to be provocative to stimulate debate and discussion about how Smart Data schemes could be designed, implemented, and monitored in the future. They are not reflective of current government policy priorities or objectives. They are not policy recommendations. They do not pass judgement on the historical, current, or future work of key stakeholders in the Smart Data ecosystem.

All of what follows is based on the thoughts of the stakeholders involved in the workshop.

Scenario 1: High public trust and low data portability

Stakeholders proposed that the rationale for high public trust in this scenario was due to consumers being aware that the data sharing ecosystem works well and that sharing their data delivers tangible value to them. This suggests an environment with high levels of transparency around how data is used and shared between organisations. Regulation is in place to control data sharing, potentially via a regulator or accredited body. If and when something does go wrong, the consumers trust that an effective safety net is in place, with proper recourse for the affected parties. The perceived ability to benefit from data sharing is also high - which means there are high levels of digital literacy and digital inclusion across society.

In spite of this high trust environment, relatively few members of the public actually share their data across and between service providers - data portability is low. The legal and regulatory landscape outlined above is more focused on risk mitigation than innovation, and so incumbent organisations are not compelled to facilitate data portability. The regulatory landscape is highly trusted, but has created high barriers to entry for Authorised Third Parties. Organisations face significant technical challenges trying to set up standardised mechanisms for sharing data. A focus on voluntary, rather than mandated, initiatives means that the standardisation of APIs for data portability has been limited, and a small number of players act as gatekeepers.

Consumers are not incentivised to switch between different service providers - partly because incumbent providers have upped their game to offer highly competitive products or services, and partly because the actual process from a consumer point of view is complex or cumbersome. Potentially, high levels of open data have fuelled an environment where data sharing does not require consumer consent for sharing.

The innovation ecosystem is relatively limited - indeed, it is one of the drivers behind this environment of low data portability. Likewise, low data portability acts as a constraint on what is possible - in other words, low data portability and low levels of innovation act in a feedback loop. There may be a “killer app” in the data custodian space which has fostered the high trust

environment - and therefore encouraged consumers not to port their data between services.[4] There may also be killer apps in the automated switching space - apps relying on open data to establish when a consumer can get a better deal with an alternative provider and automatically switch them. Again, this does not rely on portability of a consumer's data.

Limited data portability means that only limited numbers of consumers are actually using such services, but there are some alternative readings as to who these groups are. On the one hand, data portability may be a privileged service benefiting only those who have the time and resources to participate, particularly if the processes themselves are complex or cumbersome, as above. On the other hand, data portability may be limited to specific, highly regulated use cases, which are likely to be processes, such as applying for Universal Credit. In this sense, data portability could be seen as benefiting those with vulnerabilities.

The governance ecosystem in this world has some highly effective elements, which tend to be government-backed: examples from other sectors show that high levels of trust often emerge from government-backed initiatives. However, the balance between consent, accreditation, authorisation, and liability are off-kilter, leading to the constrained innovation ecosystem. For example, there may be low levels of authorisation needed to participate in the data sharing ecosystem, but the levels of liability for organisations which get something wrong are prohibitively high. Alternatively, the requirements put upon businesses seeking authorisation could be too high, restricting the number and type of organisations that can enter the ecosystem and drive innovation. This lack of balance suggests a lack of leadership - consumers are protected, but there isn't enough leadership at the data architecture level to bring about the standards that data portability needs. In other words, government leadership is in place, but not implementation or coordination leadership.

Scenario 2: High public trust in data sharing and high data portability

Stakeholders highlighted that consumers feel that they are in control of their data, contributing to high public trust in data sharing in this scenario. This is underpinned by consumers being informed about their rights and how their data is being used, as well as strong security infrastructure to protect against data breaches. There is a single regulator responsible for data sharing in Smart Data schemes, enabling consistent regulation across sectors, simplifying lines of accountability, creating guidance on fair standards, and overcoming gaps in sector-based regulators' responsibilities and oversight. This single regulator works closely with sector-based regulators, enabling regulators and other stakeholders to respond quickly to changes in the Smart Data ecosystem.

Consumers see the benefit of sharing their data, which underpins high data portability in this scenario. There is robust regulation in place to prevent misconduct on the part of data holders and/or Authorised Third Parties, and to protect consumers with vulnerabilities. Duty of care obligations have been introduced to ensure that consent by consumers is truly informed. Although the regulation mandates participation by data holders, Smart Data is underpinned by easy-to-use infrastructure - such as common data standards and APIs - which removes barriers for data holders and Authorised Third Parties. In addition, there has

[4] According to Merriam-Webster, a "killer app" is a "computer application of such great value or popularity that it assures the success of the technology with which it is associated". Merriam-Webster, *Killer app*, 2022.

been a concerted effort to increase the roll out of high speed broadband across the UK, particularly focusing on disenfranchised coastal and rural communities.

The innovation ecosystem is driven by inclusive and socially beneficial Smart Data applications, which are supported by regulators or encouraged through government-backed financing. Unmet and underserved needs are being supported through new Smart Data applications, such as financial capability apps enabling consumers with vulnerabilities to better manage their finances. Although there are concerted efforts to address digital exclusion, the majority of Smart Data services are targeted at consumers with the most financial value to service providers - as measured by either fees charged for the services or the insights gained from the data.

Scenario 3: Low public trust in data sharing and low data portability

Stakeholders highlighted that low public trust in data sharing schemes such as Smart Data is reflective of low public trust in the state in general. Changes to the data protection regime leads people to have - or perceive themselves to have - less control over their data. Consumers are sceptical about data sharing in this world: they lack visibility over what data they have shared or how companies are using their data. There is a perception that the more data consumers give, the more insights they get but there is limited capacity to prevent data providers from sharing their data with third parties. In addition, the public lack redress mechanisms for misuses of their data and data breaches, which are common and widely publicised. These factors are compounded by public debate around data sharing and data use that lacks nuance. Existing commentary is either entirely positive or negative and there are disingenuous discussions around the risks and benefits of schemes to encourage data sharing, particularly around personal data.

In this scenario, low data portability is partly due to the fact that organisations are not mandated to make data available. There are no real incentives for bigger organisations to make data portable or accessible. As a result, UK companies struggle to come together to create market-based solutions to these issues. This has knock-on effects for the wider data-sharing infrastructure: there are no technical standards - such as machine-readable formats for data sharing or APIs - and data is not required to be shared in a timely manner. Big tech companies may be the exception to low data portability in this scenario: they force consumers to share data in order to conduct consumer research. This creates an opportunity for small - and medium-sized businesses (SMEs) to support consumers in managing their data. However, consumers need to dedicate substantial hours to do so, which only appeals to a small minority.

There are low levels of innovation in this scenario as a lack of incentives to share data creates data silos. The “right” business models have not been found by providers of Smart Data-based services. As a result, there are no killer apps in this scenario. Innovation is driven by dominant players such as big tech companies who have an incumbency advantage from amassing large pools of consumer data. An alternative driver of innovation is provided in the form of layered approaches: organisations that create APIs drive innovation by enabling others to use and share data. These companies have spotted a gap in the market to provide technical standards and infrastructure that otherwise do not exist.

This is not to say that Smart Data is never used. However, it is being used in ways that work for organisations but not for individuals. The approach is largely punitive, for example: debt collectors identify when they should go to collect debt through using energy consumption data, while car financing companies can disable vehicles on a whim if drivers have not kept up payments. In addition, consumers that withdraw consent for data sharing may not be able to get a loan in future because they have withdrawn consent.

There is no governance of Smart Data schemes. The onus is on individuals to manage their data, but there are high barriers to doing so. In the absence of data protection regulations and other governance mechanisms, SMEs offering data management services rely on decentralised technologies such as the blockchain to keep track of where consumers' data is being shared. Existing personal data stores, such as SOLID, struggle to operate without a centralised governing authority. As a result, only those individuals who are both tech-savvy and privacy-conscious have the skills and time necessary to effectively manage their data.

Section 3. Analysis and reflections

Following the development of these scenarios, participants reflected on the desirable and undesirable features of the scenarios. Perhaps unsurprisingly, Scenario 2 - the high trust and high data portability scenario - was considered the most desirable scenario by all participating stakeholders, as it is a world where consumers are being adequately protected from potential harms, giving them the confidence to use the innovative new products developed by organisations. This leads to good governance, consumer convenience, and market stimulation. In their reflections on how to achieve these positive outcomes, participants emphasised the following features:

- Clear demarcation of responsibilities between regulators and/or governing bodies for Smart Data.
- Strong governance measures to mitigate conflicts of interest, data insecurity, and other such potential harms.
- Government incentives to encourage the reformatting of data and the creation of functional standards that underpin effective Smart Data use (e.g., a legislative mandate).
- Government incentivisation scheme to ensure that Smart Data for social good initiatives are included.
- Initiatives to ensure that the rights and experiences of digitally excluded or those not wanting to take part in Smart Data are protected.

In contrast, Scenario 3 - low trust and low data portability - was considered an undesirable future that was to be avoided. In this world, consumers are hesitant to share their data because of a lack of adequate regulatory protections and a distrust of governance institutions. Innovation is also being hampered by a lack of meaningful competition within the market and limited data sharing between firms. The key features of this scenario that the participants emphasised should be avoided in the future include:

- A laissez-faire approach to Smart Data, which does not provide sufficient governance to protect consumers or actively mandate and promote its use.
- Opaque use of individual data by government and organisations that further contributes to individuals' lack of control over their data.
- Market concentration of data in a select handful of organisations, which stymies innovation.
- Punitive uses of Smart Data, which benefits companies yet undermines consumer wellbeing, particularly because they are unable to opt-out of Smart Data schemes.

Reform of the wider data governance landscape in the UK will influence consumer trust and levels of data portability. On top of this wider reform, the participants had a number of practical recommendations for possible steps that could be taken to avoid the negative and help achieve the positive outcomes outlined above. These are *not* policy proposals by the CDEI or BEIS but rather provide considerations for policymakers and regulators to think about when designing future Smart Data schemes. Suggestions included:

- Creating a single regulatory body, that is well connected with sectoral regulators, that could help provide the clarity and expertise needed to effectively govern Smart Data.
- Introducing a duty of care requirement for organisations to ensure that meaningful consent is achieved.

- Encouraging government or social good enterprise initiatives to ensure the needs of consumers in vulnerable circumstances, in particular, are addressed in Smart Data schemes. Without this, commercial organisations may focus on developing products and services for more digitally-savvy, wealthier consumer segments.

Additional research would be needed to understand the viability of these stakeholder proposals. For example, a single regulatory body could provide the expertise necessary to govern Smart Data. However, it would need to navigate sectoral differences. These include - but are not limited to - differences in consumer attitudes towards different elements of their data. For example, focus group research run by BritianThinks on behalf of CDEI and BEIS with a diverse sample of 32 consumers found that the majority of participants perceived their financial data as more personal than their energy consumption data. Achieving informed consent from consumers raises similar complexities, particularly in Smart Data where it is clear that there is potential for data sharing to drive innovation in new products and services, yet it is unclear exactly what form those products will take.

Following this exercise, the CDEI undertook a number of activities for our partnership project with BEIS. This included:

- Testing the scenarios developed above with a diverse range of focus groups to understand the elements that the general public think are desirable. This included testing assumptions made by stakeholders (the government, industry, and the third sector) around consumer priorities for these schemes. This enabled the CDEI and BEIS to incorporate consumer views alongside existing industry-focused engagement.
- Developing a toolkit that builds on the findings of this exercise to guide policymakers and regulators in developing Smart Data schemes. All of the materials produced by BEIS and CDEI have been fed into an interactive toolkit that identifies the most important features of ethical and trustworthy Smart Data schemes, and offers resources to support decision-makers in implementing these schemes in practice.

Appendix - Attendees

The CDEI and BEIS are incredibly grateful to all that gave their time to be involved in this workshop. A number of attendees attended under the condition of anonymity. Some of the organisations that attendees represented included:

- The Competition and Markets Authority
- Ofcom
- The Financial Conduct Authority
- The Financial Inclusion Centre
- The Department for Digital, Culture, Media & Sport
- Expedia
- Icebreaker One
- HSBC
- NatWest
- Plaid
- TrueLayer
- Swoop Funding
- Broadband UK
- The Open Data Institute
- The University of Nottingham
- Which?