Department for Business & Trade

Smart Data: Identifying the features of ethical and trustworthy Smart Data Schemes

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) has 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. This paper summarises CDEI's research up to February 2022, and interviews with stakeholders to inform the next phases of this CDEI-BEIS programme of work.

The intended audiences of this paper are: 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.

Contents

Executive Summary

Smart Data is defined as the "secure sharing of customer data with authorised third party providers (TPPs)[1], upon the customer's request".[2] The only live Smart Data scheme is in the banking sector, however Smart Data legislation will allow schemes to be implemented in further sectors, with the authority of the Treasury and / or the relevant Secretary of State. The extension of Smart Data will give consumers more control over their data, deliver new and innovative services, create stronger competition in affected markets, and deliver better prices and choices for consumers and small businesses, including through reduced bureaucracy.[3]

Smart Data schemes have the potential to bring significant economic and social benefits for consumers and for the UK economy as a whole. Greater personal data sharing could bring around £27.8 billion per annum according to some estimates.[4] However, simply establishing Smart Data schemes may not be enough to realise their full potential benefits. Low consumer awareness or participation could be a key barrier to them realising their full benefits. Polling by the Centre for Data Ethics and Innovation (CDEI) and the Department for Business, Energy, and Industrial Strategy (BEIS) indicated that 32%of respondents surveyed thought that there would be no benefit to sharing their financial information with other organisations for Open Finance, while 33% said the same for Open Communications.[5] It is unclear whether this is due to lack of trust or scepticism around the benefits of Smart Data schemes on the part of consumers.

BEIS has been working with other government departments and regulators in the Smart Data Working Group (SDWG) to achieve interoperability, consistency, and coordination across Smart Data schemes. Sector-specific decision-makers are responsible for designing and delivering the schemes.[6] As part of this work, the CDEI and BEIS are working together to identify the features of ethical and trustworthy Smart Data schemes to inform future thinking. This paper reflects the culmination of the first phase of work, which has involved desk research, interviewing 25 stakeholders, and developing international case studies of Open Banking. It is important to note that this paper does not put forward policy proposals for how to design and implement Smart Data schemes. This paper and the recommendations presented in it are not government policy. Likewise, it is important to note that, with 25 stakeholders interviewed, this paper does not give an exhaustive understanding of the Smart Data ecosystem. This could have impacted the results of this work, and is explored in more detail in Section 1. Interview methodology and questions.

The paper suggests that Smart Data schemes should be founded on the basis of facilitating data sharing that is trustworthy, aligned with society's values, and with people's expectations of ethical behaviour, in addition to adhering to existing legal requirements. These themes align with other work the CDEI has published about ensuring trust in public sector data use.[7] This paper identifies eight features of ethical and trustworthy Smart Data schemes. This paper also explores six conditions - and corresponding interventions - that could help

[1] Following consultation with stakeholders from across the Smart Data ecosystem, the rest of 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.
[2] Smart Data Working Group, *Spring 2021 report*, June 2021.
[3] Smart Data Working Group, *Next Steps for Smart Data*, March 2021.
[4] CtrlShift, *Data mobility: The personal data portability growth opportunity for the UK economy*, 2018.
[5] The CDEI, Examining public attitudes towards Smart Data schemes, June 2022.
[6] For the purposes of this paper, decision-makers are defined as governing bodies, government departments, regulators, data holders or Authorised Third Parties/ TPPs, and consumers or consumer interest groups where relevant.
[7] The CDEI, *Addressing trust in public sector data use*, July 2020.

decision-makers to embed ethics and trust into Smart Data schemes. These include: Robust Technical Design; Appropriate, Available and Minimal Data; Clear Scope; Available Resources; Effective Governance Mechanisms; and Meaningful Engagement. This paper provides questions to prompt decision-makers' thinking, and suggests interventions that could help put these conditions into practice in these schemes.

Introduction

Smart Data is defined as the "secure sharing of customer data with authorised third party providers (TPPs), upon the customer's request".[8] In this definition, customers refer to both individual consumers as well as small- and medium-sized businesses (SMEs), particularly microbusinesses.[9] This does not restrict large businesses from participating in Smart Data schemes. Rather, it recognises that smaller businesses, particularly microbusinesses, may have less negotiating power over sharing their data in order to get better financial deals based on their personalised usage. Third party providers have previously been defined as "any authorised business or organisation that a user gives permission to access their data or with which they interact to help them navigate the market, other than their data holder(s) in that market".[10] This definition captures a wide range of organisations and could, in future, include organisations such as data intermediaries that offer personal data stores and personal information management services, among other functions.

There is currently only a live scheme in the banking sector, however the government has introduced Smart Data legislation seeking powers to enable the Secretary of State or HM Treasury to mandate industry participation in Smart Data across the economy. This was introduced into the House of Commons on 18 July 2022, as part of the DCMS Data Protection and Digital Information (DPDI) Bill.

BEIS is working with other government departments and regulators to achieve interoperability, consistency, and coordination across Smart Data schemes. This includes working closely with policy teams in DCMS that are responsible for implementing the National Data Strategy (NDS). Although these schemes may be discussed in a singular sense in the document, there could be multiple schemes within a sector in addition to multiple schemes across sectors. As a result, departments and regulators may approach designing, delivering, and implementing these schemes in different manners, especially if there are important sector-specific requirements that warrant further tailoring.

Sector-specific decision-makers are responsible for designing and delivering the schemes. 'Decision-makers' could include governing bodies, government departments, regulators, data holders, Authorised Third Parties, and consumers or consumer interest groups, where relevant. Different actors may be involved in decision-making at different phases of designing and implementing Smart Data schemes. In addition, the decision-maker groups could vary significantly depending on the sector of focus. For example, in some sectors there may be a single regulatory body or department that could assume responsibility for the scheme. Other sectors may have multiple regulatory bodies or departments. For example, the transportation sector includes bodies such as the Office for Rail and Road Regulation, Transport for London, the Department for Transport, among others. Other sectors may have no specific regulator.

The work undertaken by BEIS and the SDWG to facilitate Smart Data schemes has been identified as critical to achieving the first mission of the National Data Strategy: unlocking the value of data across the economy. The UK government is committed to creating an economy where consumers' data works for them, and innovative businesses thrive.[11] The extension of Smart Data will give consumers more control over their data, deliver new and innovative

[8] Smart Data Working Group, *Spring 2021 report*, June 2021.
[9] According to the Small Business, Enterprise, and Employment Act, microbusinesses are defined as those with less than 10 employees and it has a turnover or balance sheet total of an amount less than or equal to the small business threshold. See Legislation.gov.uk, *Small Business, Enterprise and Employment Act 2015*, 2022.
[10] Ibid.
[11] Department for Digital, Media, Culture, and Sport, *National Data Strategy*, December 2020.

services, stronger competition in affected markets, and better prices and choices for consumers and small businesses, including through reduced bureaucracy.[12]

The CDEI has been working with BEIS to identify the features of ethical and trustworthy Smart Data schemes. The aim of this paper is to pull together the CDEI's existing research and interviews with stakeholders to help inform BEIS's future thinking. The CDEI is an operational directorate within DCMS and therefore does not prescribe policy. As a result, this paper should not be read as a policy paper.

The economic and social impact of Smart Data schemes

Smart Data schemes could bring significant economic and social benefits for consumers and for the UK economy as a whole. Estimates suggest that greater personal data mobility could increase UK GDP by an estimated £27.8 billion per annum, not including the wider contribution from any digital innovations enabled.[13] Open Banking could bring yearly benefits of £12 billion for consumers and £6 billion for small- and medium-sized enterprises.[14] The benefit of Smart Data can be broken down into three broad categories:

Empowering consumers through data sharing

Smart Data schemes could create innovative goods and services that are personalised and/or customised for all consumers and could be particularly beneficial for consumers in vulnerable circumstances. The Financial Conduct Authority (FCA) defines consumers in vulnerable circumstances as customers who, due to their personal circumstances, are especially susceptible to harm, particularly when a firm is not acting with appropriate levels of care. According to this definition, vulnerability is a spectrum of risk, which can be affected by four drivers including: health (e.g., health conditions or illnesses affecting everyday life), life events (e.g., bereavement, job loss, etc.), resilience (e.g., ability to withstand financial or emotional shocks), and capability (e.g., knowledge of financial matters, digital literacy, etc.).[15]

Examples of the types of services that have been introduced through Open banking to address the needs of consumers in vulnerable circumstances include:

- **Mojo Mortgages** - combines Open Banking data with more widely used scoring methods to accurately assess what a customer can afford.[16]
- **Canopy** - uses consumer rent payments to improve credit scores.[17]
- **Tully** - provides debt rehabilitation services based on open banking data.[18]
- **Touco** and **Kalgera** - help those with mental health issues and older individuals better manage their money by building features on top of the basic aggregation proposition, such as the ability to send a notification to a trusted person if daily spending falls outside of normal patterns.[19]

Smart Data schemes could create greater consumer surplus by enabling consumers to

[12] Ibid.
[13] CtrlShift, *Data mobility: The personal data portability growth opportunity for the UK economy*, 2018.
[14] Note: The £12 billion value is an aggregate estimate derived on the basis that all UK consumers use all available Open Banking Services. Faith Reynolds and Mark Chidley, *Consumer priorities for open banking*, 2019.
[15] Financial Conduct Authority, *Finalised Guidance: FG21/1 Guidance for firms on the fair treatment of vulnerable consumers*, February 2021.
[16] Ariadne Plaitakis and Stefan Staschen, *Working Paper: Open Banking: How to design for financial inclusion*, 2020.
[17] Ibid.
[18] Ibid.
[19] Ibid.

switch to cheaper providers. Consumers on the margins of financial inclusion (defined as those without an account or access to only a basic account) are likely to pay less in fees with Open Banking - saving the equivalent of 0.8% of their income.[20] Those who are overstretched (defined as those with accounts and heavily indebted) could save the equivalent of 2.5% of their income.[21]

Improving outcomes in regulated markets
The Competition and Markets Authority (CMA) mandated Open Banking in February 2017 specifically to increase competition in the banking sector by improving consumer choice, encouraging more engagement with banking services, stimulating innovation, and unbundling services. Similar motivations could also be relevant for Smart Data schemes in other sectors, where service quality has often been poor and costly for consumers.[22]

Smart Data schemes could improve the efficiency of service provision by data holders and Authorised Third Parties. This includes, for example, automating applications for mortgages or other services that require consumer financial data, saving workers' time. Fair4AllFinance research highlights that this benefit has been particularly important for Credit Unions, who have used Open Banking to save money on the services they provide consumers.[23] One example is NestEgg, a software solution that automates loan assessments for responsible lenders, enabling these lenders to extend loans to new customers, improving service delivery, and reducing drop-offs in applications.[24] According to stakeholders, banks have also benefited from fewer drop-offs due to Open Banking removing frictions in the process of applying for new financial products and services.

Tackling pressing economic, societal, and environmental challenges
Rolling out Smart Data schemes could enable the UK to tackle pressing economic, societal, and environmental challenges. For example, sharing energy usage data could enable Authorised Third Parties to develop services for reducing consumers' carbon footprints. Other approaches could include improving the efficiency of the existing energy infrastructure network. In April 2021, IceBreaker One identified fifteen particularly promising Green Economy use cases that Smart Data could help unlock.[25]

The challenge with Smart Data schemes

Delivering Smart Data schemes may not be sufficient to realise the full benefits of these schemes. Polling conducted by the CDEI and BEIS indicated that 32% of respondents surveyed thought that there would be no benefit to sharing their financial information with other organisations for Open Finance, while 33% said the same for Open Communications.[26] Consumer participation in Open Banking services is low, but steadily increasing. According to the June 2022 Open Banking Impact Report[27], 10-11% of digitally-enabled consumers are active users of at least one open banking service, an increase from 5-6% in December 2020.[28] Research by Experian suggests that requests by consumers to share data through Open Banking-enabled services tripled during the pandemic.[29] Despite this increase, low

[20] Faith Reynolds and Mark Chidley, *Consumer priorities for open banking*, 2019.
[21] Ibid.
[22] Open Data Institute and Fingleton, *Open Banking: Preparing for lift off: Purpose, Progress, and Potential*, 2019.
[23] Fair4AllFinance, *Transforming affordable credit in the UK*, February 2020.
[24] NestEgg, 2021.
[25] IceBreaker One, How can Smart Data help unlock a Green Economy?, April 2021.
[26] The CDEI, Examining public attitudes towards Smart Data schemes, June 2022.
[27] Open Banking Implementation Entity, *The Open Banking Impact Report*, June 2022.
[28] Open Banking Implementation Entity, *The Open Banking Impact Report*, October 2021.
[29] Open Banking, *May Highlights*, May 2021.

participation in, and awareness of, Smart Data schemes will be a key barrier to them realising their full benefits.

The following sections will explore how decision-makers could address these barriers by establishing ethical and trustworthy Smart Data schemes. This paper is organised as follows: Section 1. **Interview methodology and questions** outlines the interview methodology and questions, Section 2. **Defining ethical and trustworthy Smart Data schemes** considers how to define ethical and trustworthy Smart Data schemes, outlining emerging issues and identifying eight features that could help deliver these schemes in an ethical and trustworthy manner. Section 3. **Approaches to implementing ethical and trustworthy Smart Data schemes** considers how decision-makers could implement these features across the Smart Data scheme lifecycle. Section 4. **Next phases and areas of future work** provides an overview of next phases and areas of future work. The findings in this paper have been based on desk research, engagement with twenty-five stakeholders, and international comparisons of Open Banking. The Appendix: **List of Interviewees** provides a list of interviewees.

Section 1. Interview methodology and questions

To inform this paper, the CDEI undertook 25 semi-structured interviews with academics, regulators, Authorised Third Parties/ TPPs, data holders (e.g., industry bodies), and third sector organisations such as charities between July and September 2021. This was a non-representative sample of 25 expert stakeholders. The findings of this work reflect the views of the stakeholders that participated but should not be interpreted as the views of all stakeholders. We note that there was a skew in representation towards those with interest and experience in the financial sector. More details on the organisations interviewed are given in the appendix.

Semi-structured interviews were selected because they provide interviewees with the scope to guide us towards the types of ethical questions that concerned them most, whilst ensuring that conversations stayed on track.

To analyse our interview data, we employed two cycle coding. This entailed inductively analysing our data to develop emerging themes, of which there were around 50, and then organising and grouping these through a second cycle of coding (outlined in the *Table 1* below). Our aim from this coding exercise was to identify the key ethical and governance challenges that interviewees considered to be associated with Smart Data.

Table 1: Interview themes

Meta-theme	Theme	Description
Context	Different countries, different rationales	Open Banking has been introduced as a demand side remedy in the UK but has been introduced for different reasons and governed in different ways elsewhere.
	Context is key	Smart Data schemes have a number of similarities but also different levels of risk and context specificity.
	It takes time to see benefits	Open Banking has achieved a lot, including for social good, but it will take time for all the benefits of this and other Smart Data schemes to materialise.
	Big Tech could shake things up.	Big Tech has the data and infrastructure to be impactful in various Smart Data markets.
Public trust	Public awareness campaigns are contentious	There is a lack of consensus over whether public awareness campaigns and/or kitemarks would be beneficial.
	There just isn't that killer app yet	People would be willing to use Smart Data if there was a better (marketed) app that fully served their needs (e.g. write functions).
	Engagement and trust isn't uniform	The level of consumer engagement and trust in Smart Data will differ depending on their specific contextual circumstances. This can be due to personal experiences and/ or structural issues

	Complexity, communication and consent.	Smart Data is complex which raises challenges for communicating with consumers and ensuring that consent is present.
Governance	Mandate Smart Data	Smart Data needs to be mandated to work effectively for the customers' interests.
	Consent could be better	How or whom people should give consent to needs to be thought about more for Smart Data.
	Multistakeholderism, not techno-solutionism	Open Banking was viewed as a technical solution, meaning sufficient thought has not been given to building trust and serving groups equally. A multistakeholder, inclusive conversation that is ongoing is needed for the long-term success of Smart Data.
	Privacy enhancing technologies	PETs have the opportunity to lessen the collection and analysis of personal data which could improve public trust.
	Accreditation and liability	Ensuring that clear accreditation and liability mechanisms are in place is necessary for building public trust in the Smart Data ecosystem.
	Data quality and bias	Bias can creep in if data quality, including representativeness is inadequate.

Section 2. Defining ethical and trustworthy Smart Data schemes

The CDEI's stakeholder engagement work highlighted that there are differing approaches to defining data ethics in relation to Smart Data schemes. Some stakeholders argued that data ethics should primarily focus on data rights, establishing a minimum bar that addresses the needs of consumers in vulnerable circumstances, whilst extending the benefits of these protections to all consumers. The CDEI suggests a more holistic approach:

> *Smart Data schemes should be founded on the basis of facilitating data sharing that is trustworthy, aligned with society's values, and with people's expectations of ethical behaviour, in addition to adhering to existing legal requirements.*[30]

The benefits of ethical and trustworthy Smart Data schemes

We have identified three main benefits to establishing ethical and trustworthy Smart Data schemes: fostering legitimacy, encouraging consumers to participate, and driving innovation in Smart Data schemes.

Fostering legitimacy in data sharing activities by encouraging consumers to participate

Previous work by the CDEI highlights that the legitimacy and sustainability of data sharing activities requires more than the legal frameworks upon which organisations are allowed to share data. Data sharing activities require broader public consent and a level of public trust, in addition to legal frameworks.[31] Building this public trust could happen in part by de-risking business or organisational engagement in Smart Data schemes, making that engagement more attractive for already trusted brands, thereby raising consumer engagement with these schemes. Some stakeholders the CDEI interviewed highlighted that **consumers are more likely to trust prominent brands**, such as their banks or credit unions, and participate in Open Banking schemes if their financial institution recommends they do so. For example, one stakeholder's user experience research demonstrated higher conversion to Open Banking integration through their partnering bank's website than through the authorised service's website alone. As a result, **making it more appealing for large data holders to participate in Smart Data schemes** could bring in more participation by consumers and Authorised Third Parties.

Driving innovation in Smart Data schemes

Although trust or trustworthiness are abstract concepts, the knock-on effects of losing trust can undermine innovation and restrict an organisation's ability to adopt new tools or undertake their existing work. For example, the General Practice Data for Planning and Research (GPDPR) programme intended to gather public healthcare data to find better treatments and improve patient outcomes. However, media and public concern surrounding the use of NHS data and the limited time to inform patients about their rights to opt-out led to implementation being postponed.[32] By contrast, **if trust and trustworthiness are established within Smart Data schemes, there should be greater scope for future innovation**.

Replacing current practices that may be unethical and untrustworthy

[30] The Centre for Data Ethics and Innovation, *Addressing trust in public sector data use*, July 2020.
[31] Ibid.
[32] NHS Digital, *General Practice Data for Planning and Research (GPDPR)*, 2021.

Open Banking-enabled services offer an alternative to practices such as screen scraping, in which organisations construct "an agent to download, parse, and organise data from the web in an automated manner".[33] This practice is also known as web scraping, web data extraction, or web data mining. There is potential for this practice to, in future, infringe on consumers' privacy: users often need to give their accounts and login credentials in order for these services to work. Given that users frequently use the same login details for multiple online accounts, this could compromise their overall online safety. In addition, sharing these details could be a direct breach of the terms of service a consumer has with their bank or other providers.[34]

There are limited mechanisms for holding scraping services accountable for misuses of the data they collect. For example, if a screen scraper conducts an unauthorised transaction on a user's account without their consent, banks may not be liable - or hold themselves liable - for the transaction.[35] There is no accreditation of screen scraping providers and no regulation or restrictions on how they use the data they collect. Addressing these limitations is one way in which Smart Data schemes offer a more ethical, trustworthy alternative to existing practices.

Emerging ethical and trust-related issues in Smart Data Schemes

The CDEI's desk research and engagement with stakeholders highlighted ethical and trust-related risks that could be addressed in the design, implementation, and ongoing delivery of Smart Data schemes. *Table 1: Principles and potential risks* provides a high-level summary of these findings, which are discussed in more detail below.

Table 2: Principles and potential risks

Principle	Potential risks
Fairness	Consumer groups in vulnerable circumstances pay higher prices for the same goods and services.
	Consumers in vulnerable circumstances lack access to goods and services relevant for their needs.
	Algorithmic decision-making is unfairly biased against consumers in vulnerable circumstances.
Privacy	Consumer data is misused to exploit certain consumer groups.
	Consumer data infers details about consumers that are classified as special category data.
Consent	Consumers are ill-informed about how their data is being used and shared.
Accountability	Complexity of roles and responsibilities in Smart Data schemes creates liability challenges, and raises questions around how

[33] Open Data Institute and Fingleton Associates, *Data sharing and Open Data for banks: A report for HM Treasury and Cabinet Office*, September 2014.
[34] Han-Wei Liu, "Two decades of laws and practice around screen scraping in the common law world and its Open Banking watershed moment", *Washington International Law Journal* Vol 30. No. 1, 2020; Financial Consumer Agency of Canada, *Open Banking*, 2021.
[35] Ibid.

	consumers can seek redress.
Security	Aggregation of consumer data creates targets for hostile actors.
	Potential for consumer data to be used, shared, and stored insecurely.

Fairness

Consumers in vulnerable circumstances pay higher prices for the same goods and services. This could be exacerbated in Smart Data schemes if the onus remains on consumers to actively seek out better deals. Polling by the CDEI and BEIS suggests that 42% of higher-income households have switched services, such as switching banks, compared to 24% of lower-income households.[36] Existing research suggests that lower-income households pay a poverty premium of £233 per year more on energy on average than higher-income households due to not switching fuel suppliers.[37]

Consumers in vulnerable circumstances lack access to goods and services relevant for their needs. Consumers in vulnerable circumstances may have unique needs that are not addressed by services developed for more financially or digitally-savvy consumers. For example, existing research highlights that although many lower-income households could benefit financially from switching to cheaper fuel tariffs, many have valid reasons for not switching. In energy, pre-pay meters may be more expensive yet they provide households with predictability - and therefore more control - over their finances.[38] In addition, these consumer segments may be less financially lucrative for data holders and Authorised Third Parties, reducing the incentive to develop services for them.

Algorithmic decision-making is unfairly biased against consumers in vulnerable circumstances. Numerous studies have shown that training AI systems using data that is not adequately representative can lead to biased and unfair outcomes.[39] The underutilisation of Smart Data applications by some consumers with vulnerabilities risks discriminatory outcomes for consumers within that group that do use said applications. Polling by the CDEI and BEIS - in addition to interviews with stakeholders - suggests that consumers do not fully understand the implications of bias in algorithmic decision-making. Although 15% of respondents were aware of "computers making wrong decisions because the data they are using is biased against certain types of groups and types of people", they were more likely to be aware of issues associated with internet-connected devices in homes (53%), targeted advertising online (51%), and the use of facial recognition technology (48%) in relation to data collection and use.[40]

Privacy

Consumer data is misused to exploit certain consumer groups. Consumers whose financial data demonstrates they regularly dip into their overdraft or gamble online may be identified as more risky candidates for financial loans, thereby raising potential interest rates for these consumers. 38% of respondents expressed concerns that there would not be sufficient safeguards in place to ensure companies will serve their interests in the financial services

[36] The CDEI, Examining public attitudes towards Smart Data schemes, June 2022.
[37] Sara Davies, Andrea Finney, and Yvette Hartfree, *Paying to be poor: Uncovering the scale and nature of the poverty premium*, November 2016; University of Bristol, *The poverty premium: When low-income households pay more for essential goods and services*, November 2016.
[38] Ibid.
[39] The CDEI, *Interim report: Review into bias in algorithmic decision-making*, July 25, 2019; Nicol Turner Lee, Paul Resnik, and Genie Barton, *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms*, The Brookings Institute, May 22, 2019.
[40] The CDEI, Examining public attitudes towards Smart Data schemes, June 2022.

sector, while 35% of respondents expressed the same for the telecommunications sector.[41]

Consumer data leads to inference of details about consumers that are classified as special category data. Special category data includes personal data revealing racial or ethnic origins, health status, political opinions, sexual orientation, among others.[42] There is a risk for consumers that these categories could be inferred, intentionally or unintentionally from data such as subscriptions, spending patterns, energy usage at specific times of the day or year.[43] There are strict regulations around special category data and requirements for processing it in UK GDPR, which can include inferences made about individuals based on their personal data, depending on how certain the inference is and whether it is being deliberately drawn.[44]

Consent

Consumers are ill-informed about how their data is being used and shared. There are questions as to whether consent can meaningfully be given by consumers given the complexity of some Smart Data applications, particularly for consumers in vulnerable circumstances and when algorithmic processing is used. Other concerns include how frequently consent should be sought after initial consent has been given, and what constitutes a material change in use case that requires seeking further consent.[45] 24% of respondents (from a sample of 2,000) said they wanted clear and accessible explanations of who can access the data and for what purpose for Open Finance and for Open Communications.[46]

Liability

Complexity of roles and responsibilities in Smart Data schemes creates liability challenges. One challenge in Open Banking is determining who is liable for data security in the event that an Authorised Third Party ceases trading - currently, banks and regulators could be held liable for the loss of consumer data.

Security

Aggregation of consumer data creates targets for hostile actors. Smart Data schemes could create new large datasets, with highly personal consumer data, that become valuable targets for hostile actors. Consumers could be adversely affected by these actors gaining access to these datasets.

Potential for consumer data to be used and shared insecurely. Smart Data schemes will by definition open up new interfaces to existing large personal datasets, increasing the attack surface for cyber attacks.

Consumer concerns about security are also an issue. Polling by CDEI and BEIS suggests that 40% of respondents expressed concerns that their data will not be shared and used securely in the financial services sector, while 43% expressed the same in telecoms.[47]

Many of these risks are not unique to Smart Data schemes. For example, the financial sector

---

[41] Ibid.
[42] Information Commissioner's Office (ICO), *Special category data*, 2021.
[43] Flavio D. Garcia and Bart Jacobs, "Privacy-friendly energy-metering via homomorphic encryption", *Security and Trust Management*, 2010, pp. 226-238.
[44] Information Commissioner's Office (ICO), *Special category data*, 2021.
[45] Miles Cheetham, Faith Reynolds, Sharon Cunliffe, Gavin Starks, *BEIS: Smart Data: Report: Consent*, March 2020; Behavioural Insights Team, Doteveryone, and Centre for Data Ethics and Innovation, *Active online choices: Designing to empower users*, November 2020.
[46] The CDEI, Examining public attitudes towards Smart Data schemes, June 2022.
[47] Ibid.

already segments consumer data to minimise their exposure to risk and target specific products to certain consumer segments. Screen scraping and other data sharing initiatives also face issues around who has access to the data and the safety and security of personal data. Designing and implementing Smart Data schemes could help to address many of these risks. However, these schemes should also address specific risks that could arise as a result of their design and implementation, such as insufficient services designed for consumers in vulnerable circumstances and consumers' concerns around who has access to their data within these schemes.

The emerging features of ethical and trustworthy Smart Data schemes

This section gives eight features for ethical and trustworthy Smart Data schemes. These features have been drawn from the existing work on trust by the CDEI and tailored to the Smart Data context through incorporating the perspectives of stakeholders and desk research.
These features are:

1. **Clear lines of *accountability* in and across Smart Data schemes.** This should include openly and clearly stating the roles and responsibilities of governing bodies in and across schemes.
2. ***A clear purpose* for the scheme.** This could include, for example, improving value for consumers by introducing more competition into a market, or introducing wholly new services into a market.
3. **Ensuring *fair outcomes* for consumers, without creating unnecessary restrictions to innovation.** Firms and governing bodies should clearly articulate the value of sharing consumer data in exchange for better services on the basis of fairness, and in ways that appreciate the potential for future innovation.
4. **Respecting and protecting consumers' *privacy* in the design and delivery of Smart Data schemes.** Unique approaches in schemes could be sought when the sensitivity of the data being shared, or the effect of its use on consumer trust, is particularly significant.
5. **Ensuring Smart Data systems *function as intended* so that they do not inflict physical or mental harm.** This could include informing consumers of their right to withdraw consent or encouraging firms and governing bodies to seek extra ways to inform consumers about what they are agreeing to when sharing their data in order to gain consent.
6. **Stringent data *security controls*.** Smart Data schemes could include materially improving the existing data sharing landscape by protecting individuals' privacy, while also facilitating better data sharing.
7. **Delivering clear *benefits to society* as a whole.** Smart Data schemes could unlock financial benefits worth billions annually for consumers, increase competition in markets with poor consumer outcomes, as well as lead to innovations that tackle economic, social, and environmental challenges.
8. **Being *transparent* and open to scrutiny by participants and other legitimate sources.** The use and sharing of data by Authorised Third Parties should be open to scrutiny by consumers, and appropriate behaviour should be open to redress.

Section 3. Approaches to implementing ethical and trustworthy Smart Data schemes

This section considers approaches for implementing Smart Data schemes in an ethical and trustworthy manner, leading on from the features identified in Section 2. **Defining ethical and trustworthy Smart Data schemes**. It provides a non-exhaustive set of questions to prompt decision-makers' thinking, suggests possible interventions that could help put these conditions into practice, and summarises case studies of Open Banking in other geographies. Ethical and trust-related concerns could arise across all of these conditions and throughout the design, development, and implementation of Smart Data schemes.

In total, this paper considers six conditions: Robust Technical Design; Appropriate, Available and Minimal Data; Clear Scope; Available Resources; Effective Governance Mechanisms; and Meaningful Engagement. Some of these conditions may require more attention and focus in a given Smart Data scheme. For example, stakeholders highlighted that Open Banking was implemented as a technical solution to address consumer outcomes (and therefore prioritised Robust Technical Design), but more focus could have been placed on engaging with consumers (e.g., Meaningful Engagement). The ways in which each of these six conditions could be fulfilled may vary depending upon the sector, the specific scheme being introduced, and the decision-maker involved. There may be unique challenges in expanding these schemes across sectors as well. While the questions below have been attributed to specific decision-makers, they may be relevant for more than one decision-maker, depending on the specifics of the scheme.

Robust Technical Design

Smart Data schemes should be designed with the following in mind:

- The *functional* design: how the data-sharing system may behave in relation to outside agents;
- The *technical* design: how this functionality may be implemented in code;

All Stakeholders interviewed highlighted that privacy-by-design will be important for building public trust in Smart Data schemes. As a result, the systems underpinning these schemes should be designed such that they are privacy-preserving, and the rights of individuals around their personal data are respected. This includes being measurably secure and resistant to being compromised. Data sharing systems - such as the APIs provided by a data holder - should have end-to-end auditability. This should include encouraging those maintaining the systems to be able to validate that they are behaving as intended and as expected at all points across their lifecycle. Considering the following questions could enable decision-makers to ensure Robust Technical Design in Smart Data schemes:

Decision-maker:	**Questions for decision-makers to consider:**
Scheme Governing Body(s)	• Is there a need for/would this scheme benefit from centrally managed infrastructure? (see below: Box 1, SGFinDex) • What technical standards are necessary to ensure the safe and smooth transfer of data between organisations? (e.g. Should there be standards ensuring that data is always formatted in the same way, as with Open Banking?) • What level of security (of both the data and the system more

	generally) will be required, particularly when considering the level of sensitivity of the data involved? ● What needs to be done to ensure these standards remain sufficient and relevant in the future? How will the standards be monitored and, if necessary, altered?
Regulator(s)	● What systems could be put in place to ensure that participants in the scheme adhere to its standards and security features? ● What is the appropriate level of monitoring, at the scheme, architecture and consumer outcome levels, to ensure that a scheme is operating as intended?
Data Holders/Authorised Third Parties	● Do you adhere to all of the standards and security features mandated and recommended by the governing body? ● What protocols could be required for testing, validating, verifying and monitoring the safety and reliability of your data sharing architecture? (see below: Box 1, SGFinDex) ● Should there be structures in place to test that when systems face adversarial conditions they continue to: ○ Maintain integrity and remain functional and accessible? ○ Perform reliably and accurately? ○ Keep private information secure?
Cross-sector considerations:	● What could be done to ensure interoperability between sectors? ○ Should there be cross-scheme technical standards? ● Who should be responsible for setting technical and security standards across schemes? ○ Who should be responsible for ensuring these standards remain sufficient and relevant in the future, and altering these standards if necessary? ● Would centrally managed infrastructure aid collaboration and innovation across schemes? (See below: Box 1: SGFinDex) ● Who should be responsible for ensuring participants are adhering to standards and security features? (e.g. Should it be regulators from within each sector, or should one regulator bear overall responsibility?)

Possible interventions

The following possible interventions could be used to test the robustness of Smart Data systems and ensure that they remain safe and secure, even under hostile or adversarial conditions. These include:

- **Use of privacy preserving/enhancing technologies:** see below: Box 2: PETs;
- **Penetration testing**: authorised, simulated cyberattacks performed to evaluate the security of a system;
- **Bug bounty programs**: organisations could reward individuals that report bugs, particularly those related to security exploits and vulnerabilities.

Box 1: SGFinDex

The Singapore Financial Data Exchange (SGFinDex), launched in 2020, is a public digital infrastructure that uses Singapore's national digital identity (SingPass), and a centrally managed online consent system, to allow people to access financial information held on them by government agencies and participating financial institutions. SGFinDex is a collaboration between the public sector, The Association of Banks in Singapore, and seven participating banks. Data available through SGFinDex includes government data (such as social security, tax, and Medisave data), banking data, and investment data. There are plans to onboard more financial institutions to SGFinDex including insurers and the Singapore Exchange (SGX) Central Depository (CDP).[48] Users are able to access all of this data through one central portal.

This centrally run infrastructure differs from that of Open Banking in the UK, where data holders and Authorised Third Parties/ TPPs share directly between one another (see below: Exhibits 1 and 2, respectively). SGFinDex is highly reliant on the existence of SingPass, of which there is currently no equivalent in the UK. However, a national digital ID is currently in development in the UK.[49]

Some of the stakeholders the CDEI engaged with, primarily those representing data holders, suggested that one advantage of SGFinDex structure is that it is cheaper to establish and run for the institutions involved as it requires the creation and maintenance of only one set of data sharing architecture. However, a potential downside to this type of structure is that it provides a targetable single point of failure or vulnerability. In addition to this, SGFinDex is designed around specific use cases, and it provides no opportunity to share data points beyond those specified by the Monetary Authority of Singapore (MAS). As such, stakeholders flagged that this structure may not be as well suited to fostering innovative use cases as Open Banking is in the UK.

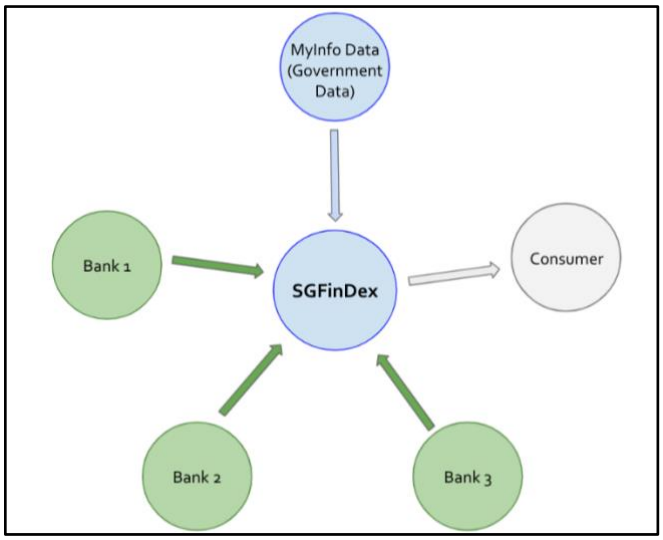

Exhibit 1: Data sharing in Singapore through SGFinDex. Green represents banks/data holders, blue represents government data holders/government-run institutions, grey represents consumers.

[48] The Association of Banks in Singapore, *Singapore Financial Data Exchange (SGFinDex)*, 2021.
[49] Matt Warman MP and DCMS, *Next steps in plans to govern use of digital identities revealed*, August 2, 2021.

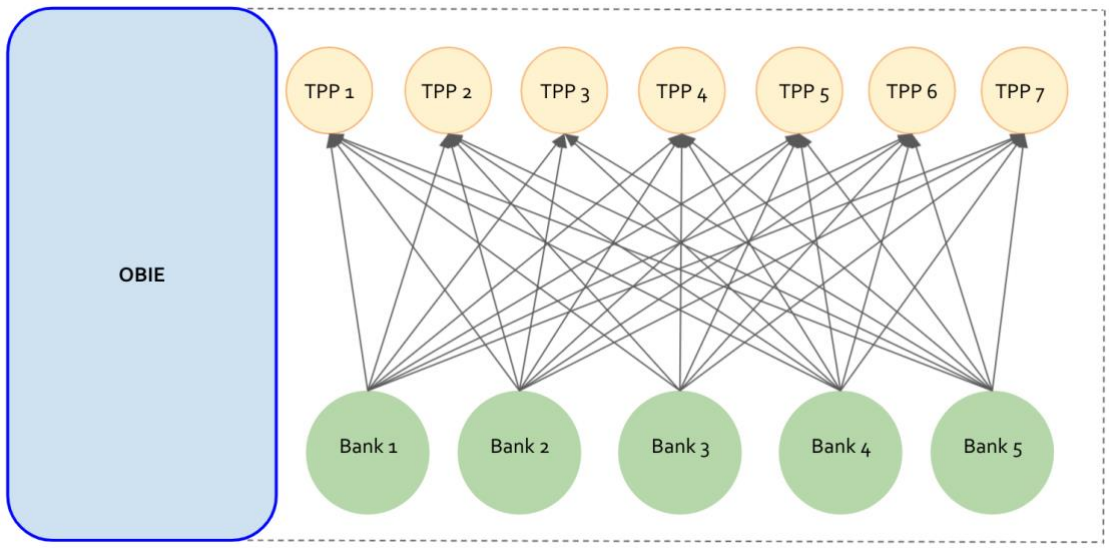

Exhibit 2: Data sharing in the UK under Open Banking. Green represents banks/data holders, yellow represents Authorised Third Parties/ TPPs, blue represents the governing body.

In the example given in Exhibit 2, each individual data holder is responsible for creating and maintaining the data sharing architecture.

Box 2: Privacy-enhancing technologies (PETs)

A key challenge of providing access to sensitive data is balancing privacy and transparency. In order to make use of data, it needs to be accessible to the data user, which necessarily compromises data subjects' privacy to some degree. A set of emerging privacy-enhancing technologies (PETs) are beginning to shift this trade off so that high levels of both privacy and transparency can be achieved. These technologies can enable an individual or organisation to answer questions using data they cannot see.[50] The CDEI's PETs Adoption Guide explores how PETs can be applied in practice.[51]

Smart Data schemes could leverage PETs to grant Authorised Third Parties access to sensitive consumer data in a more privacy-focused way. For example, financial institutions are currently using PETs to collectively analyse sensitive transactional data from financial institutions to identify fraud coordinated across these institutions.[52] These types of technologies could be useful when consumer data is aggregated, and used to develop and design new services.

Appropriate, Available, & Minimal Data

Smart Data schemes should be designed such that the data needed to enable consumers to benefit is made shareable (both personal data and data more generally, such as data around service provision). There are existing legal requirements, set out in UK GDPR, around ensuring the minimum amount of personal data required is shared in each use case.[53] The

[50] *OpenMined*, 2021.
[51] Centre for Data Ethics and Innovation, *PETs Adoption Guide*, 2021.
[52] The Future of Financial Intelligence Sharing, *The FFIS Enhancing Technology (PET) project*, January 2021.
[53] Information Commissioner's Office, *Principle (c): Data minimisation*, 2021.

Information Commissioner's Office (ICO) is tasked with regulating information rights and data protection legislation, including the UK GDPR. Readers of this document might also wish to refer to the ICO's Guide to Data Protection and data sharing information hub. Compliance with data protection law is an important practice that will help engender and maintain trust. Beyond legal requirements, ensuring this data is relevant, recent, and accurate can help minimise the risks associated with data loss, help build consumer trust in a service, and reduce the costs associated with data storing and sharing.

The above is not intended to restrict organisations from requesting substantial amounts of data for specific use cases, or reduce the overall capacity for data sharing in the economy. Rather, it encourages organisations to consider the data they require for their clearly articulated purpose and ensure appropriate safeguards are in place to protect consumers. For example, for a use case such as conducting research on the financial wellbeing of consumers during the COVID-19 pandemic, the minimum level of data required will likely be relatively high, encompassing financial data, demographics, healthcare data, and employment data. For a use case such as comparing energy tariffs, the minimum amount of data will likely be relatively low, likely encompassing a consumer's energy usage alone.

Providing meaningful options to consumers on the amount and type of data they wish to share, and the use cases for which they are sharing the data will be important. This might include allowing consumers to specify how long their data is used and stored for. However, determining how long data is needed is a complex question. Stakeholders highlighted a number of legitimate reasons that an organisation may wish to retain data for extended periods of time (for example, to monitor the outcomes of their services). In such cases, organisations might consider anonymising personal data, as anonymised data falls outside of the scope of data protection legislation in the UK.[54] If data will be aggregated and used for other purposes by either data holders or Authorised Third Parties, organisations should try to ensure that this data is representative. If this is not feasible, data holders or Authorised Third Parties might consider how they can mitigate the issues that may arise from uneven representation, such as addressing unfair outcomes for consumers in vulnerable circumstances. Considering the following questions could help decision-makers implement Appropriate, Available, and Minimal Data in Smart Data schemes:

Decision-maker:	**Questions for decision-makers to consider:**
Scheme Governing Body	● What elements of a consumer's data should data holders be required to share to maximise the potential benefits of a Smart Data scheme? ● How long should organisations be allowed to retain shared data? ● Should there be data exchange reciprocity (the two-way sharing of data between participants in a scheme)? (see below: Box 3, Data Exchange Reciprocity) ● Is it appropriate to allow data holders and Authorised Third Parties to aggregate data shared by consumers for purposes beyond those directly relevant to the consumer's use case? ○ Have consumers provided consent to data holders and Authorised Third Parties using data beyond specific use cases? Are they able to opt out? ○ Are further uses of the data justified by serving the

[54] Note: The government's *Data: A New Direction* consultation looks to address the issue of data minimisation and anonymisation in more detail, namely around clarifying what constitutes anonymised data. Department for Digital, Culture, Media and Sport (DCMS), *Data: A New Direction*, 2021, pp.44-47.

	public interest/societal well-being? ○ If data is being aggregated and used as parts of broader datasets, are these datasets representative of the population under consideration? If not, is there transparency over what the aggregated dataset does and does not contain? ○ What could the potential impacts of a misrepresentative dataset be, particularly for consumers in vulnerable circumstances? ○ Where datasets are being aggregated, are appropriate steps being taken to protect individuals' data/anonymity? ○ Could technologies like PETs be used? (See above: Box 2, PETs) ● What measures could be put in place to ensure that information about a person, inferred from data about another consumer, is not shared if they have not consented to sharing it? (For example, stakeholders highlighted that bank statements from one consumer might contain information about financial transfers to another consumer).
Regulator(s)	● What systems should be put in place to ensure that participants in the scheme adhere to the data sharing standards set by the governing body?
Data Holders/ Authorised Third Parties	● What systems or processes could be put in place to establish what pieces of a consumer's data are necessary for a particular use case? ● Have you informed customers about the extent of the data to be collected and shared in an understandable and digestible way? ● Have consumers consented to data sharing and processing? ● How could you ensure that you only share the data about a consumer needed for a particular use case? ● What systems do you have to ensure that data is held for the minimum amount of time necessary? ● If you are aggregating data for further use cases (e.g., training predictive models), have you considered if the datasets are representative of the population as a whole? ○ If they are not, how are you addressing the possibility of unfair outcomes against portions of the population, such as consumers in vulnerable circumstances?
Cross-sector considerations:	● What standardised formats for data might be necessary to ensure interoperability across sectors? ● What measures could be put in place to ensure that only appropriate data points are used for each specific use case? (e.g., is it acceptable for a consumer's social media data to be used when determining if they should qualify for a loan?) ● Could aggregating data across sectors enable the emergence of data monopolies or place too much power in the hands of specific organisations? (see below: Box 3: Data Exchange Reciprocity) ● Who should be responsible for ensuring that participants in

	the scheme adhere to the data sharing standards set by the governing body?

Box 3: Data exchange reciprocity

In 2017, Australia became the first country to introduce data reciprocity in Open Banking with its Consumer Data Right (CDR).[55] Previous Open Banking frameworks, including the UK Open Banking Standard and the EU Second Payment Services Directive (PSD2), make payments information accessible to Authorised Third Parties/ TPPs but have no requirement for Authorised Third Parties/ TPPs to provide customer data back to banks. Whilst the UK GDPR's right to data portability could, in theory, be leveraged to facilitate reciprocity, in practice, the lack of obligation to respond in a timely manner means that GDPR cannot be used in this way. In the UK, firms have up to thirty days to respond to a data portability request.[56]

Critics argue the lack of data exchange reciprocity amounts to an unfair and regulatory-driven competitive disadvantage for data holders. Some argue that it creates an unfair, one-sided relationship between the banks and Authorised Third Parties/ TPPs. Others go further and argue that, as big tech companies can act as Authorised Third Parties/ TPPs, it further centralises data in the hands of a few select technology companies who have access to vast datasets they are not required to share and is, as a result, anti-competitive.[57] It is because of this issue of competition that lawmakers in Australia's 2018 Open Banking review argue that ensuring reciprocity will lead to a "more vibrant and creative" system and promote greater competition.[58] It is likely that the same arguments, about centralising control in the hands of big tech players, could be made about other sectors if Smart Data schemes are implemented without data exchange reciprocity. Data exchange reciprocity also exists in a number of other markets outside of Australia, including Brazil and India.

Possible interventions

The following are potential interventions decision-makers could take around system design, specifically looking at data, namely its quality - encompassing representativeness, relevance, recency, and accuracy - and how it is stored and shared. These are interventions that could be suggested or mandated by a scheme's governing body or department, or could be adopted independently by participants in a scheme. These types of interventions include:

- **Data minimisation practices**: organisations can minimise the amount of data collected and the time data is retained.[59]
- **Data integrity techniques**: these seek to ensure that data remains accurate and consistent across its entire life cycle. This is a broad term that could include things such as: backing up data, encrypting data, and using access controls to data;
- **Data profiling**: this is the process of examining, analysing and creating summaries of data and datasets. This could enable organisations to have a more holistic view of the data they hold, and in particular assess whether it is representative of the population as a whole.

[55] Department of the Treasury, the Australian Government, *Consumer Data Right Overview*, September 2019.
[56] ICO, *Right to data portability*, accessed September 27, 2021.
[57] Amy Borrett, "How Australia is challenging the UK on open banking", *The New Statesman*, April 14, 2021.
[58] Department of the Treasury, the Australian Government, *Consumer Data Right*, May 9, 2018.
[59] Access Now, Data Minimization: Key to Protecting Privacy and Reducing Harm, May, 2021; Information Commissioner's Office, Principle (c): Data minimisation, 2021.

Clear Scope

Decision-makers should consider, clarify, and communicate the intended outcomes of introducing a new Smart Data scheme, and from this, the scope of the scheme. Establishing a scope, and corresponding boundaries, will help prevent unintended negative outcomes emerging as a result of the scheme, whilst still allowing as much innovation as possible. Decision makers may find it useful to consider the following when developing the scope of a scheme:

- the intended use cases of the scheme,
- the data that would be needed,
- potential use cases emerging from the scheme, and
- opportunities created by cross-sector sharing.

Decision makers might consider how the scope of a scheme can be established such that it will provide beneficial outcomes for individuals, societies, and the planet. As schemes will evolve through time, decision-makers might consider how they can design oversight mechanisms such that they are able to alter the scope of a scheme as appropriate.

Where a scheme is intended to foster growth of new market offerings, decision-makers will need to consider how to prevent unintended negative outcomes from emerging. As mentioned earlier, some stakeholders argued that one unintended negative consequence of Open Banking in the UK is that it could give more data, and therefore more power, to a few select technological players (see above: Box 3: Data Exchange Reciprocity).

Those designing a scheme may wish to establish processes to avoid mission creep and have an understanding of the potential impacts and consequences if participants do not adhere to boundaries, intentionally or otherwise. However, they should be mindful that the boundaries placed around these schemes should not stand unnecessarily in the way of potential future innovations and opportunities. Considering the following questions could help decision-makers implement Clear Scope in Smart Data schemes:

Decision-maker:	**Questions for decision-makers to consider:**
Government Department(s)	● What is the purpose for introducing this Smart Data scheme? ○ Would the outcomes from a Smart Data scheme be measurably better than existing options/offer entirely new market offerings? ○ Does the societal benefit from introducing the scheme outweigh any potential costs?
Regulator(s)	● What are the overarching goals of the Smart Data scheme as stated by relevant legislation? ○ What knock-on implications does this have for determining boundaries of the scheme? (See below: Box 4: Data sovereignty in Australia). ● Have you assessed the trade offs between creating firmer boundaries as to how data can be used/what data is used, and the scope for more innovation? ○ Where should the boundary lie such that it maximises potential for innovation whilst limiting potential harms?
Scheme Governing Body	● What boundaries exist to establish when it is permissible for data to be shared and how it is used? ○ Have these boundaries been adequately explained to

	consumers? ○ Have you established how data holders and/or Authorised Third Parties should communicate these boundaries to consumers? (For example: this should be in simple English, easy to understand, and in a standardised format).
Data Holders/ Authorised Third Parties	● Have you adequately communicated to consumers what elements of their data are being used, how it is used, and the limits to this use? (For example, this should be in simple English, easy to understand, and a standardised format). ● Do you have procedures in place to terminate data sharing if boundaries are breached?
Cross-sector considerations:	● What are the potential consequences that could arise from combining data from different Smart Data schemes/sectors? ○ Is keeping data siloed between sectors desirable? ■ What additional risks could be created by combining data? ■ Are there benefits to keeping data siloed? ■ What additional opportunities could be gained through combining data? ● When should it be appropriate for an organisation to access a consumer's data from another sector? ● Could encouraging data sharing between sectors increase the risk of negative outcomes for consumers in vulnerable circumstances? ● Are there risks that mandating cross-sector data sharing could centralise power in the hands of a few organisations?

Box 4: Data sovereignty in Australia

Whilst many of the reasons for introducing Smart Data schemes in different countries are shared, the emphasis and weight given to each reason can have significant impacts on how the schemes are developed and implemented. For example, the Consumer Data Right (CDR) in Australia came as part of the Australian Government's commitment to give Australians greater control over their data.[60] By contrast, in the UK, HM Treasury announced its commitment to introducing Open Banking as a means of increasing competition in the 2015 Budget.[61] This is not to say that many of the expected benefits of the schemes weren't shared, for example the Australian government argued the CDR would improve access to goods and services, improve consumer outcomes, and in fact even identified that the CDR would improve competition in the marketplace.

This difference is important because it has a significant knock-on impact. In Australia, the CDR implies that consumers should own and control their data regardless of sector. This means that the CDR will necessarily be implemented in sectors outside of the banking and financial sector. The CDR will first be applied to the banking sector, then the energy and telecommunications sector. The decision to first implement the CDR in banking is simply a practical one. By contrast, Open Banking was introduced in the UK to induce competition

[60] Department of the Treasury, the Australian Government, *Open Banking Guidelines for Open Data Participants*, July 2018.
[61] Ibid.

specifically in the banking sector.[62] The government has introduced Smart Data legislation seeking powers to enable the Secretary of State or HM Treasury to mandate industry participation in Smart Data across the economy. While legislation will provide enabling powers for future Smart Data schemes to be established in other sectors, any future scheme will require secondary legislation before it could be implemented. The sector specific approach in the UK means that decision-makers will have to consider the specific issues they seek to address when introducing these schemes.

Possible interventions

The following interventions could be implemented by decision-makers, particularly Authorised Third Parties and data holders, either by mandate or voluntarily, as a way of preventing and mitigating potential downstream consequences or harmful effects:

- **Impact and risk assessments**: these are assessments that look at parts of a system that are vulnerable to attack, as well as the impact and risks associated with the intended use of a given technology.
- **Data governance frameworks**: these are rules surrounding data use in an organisation and establishing who has authority and control over data assets.
- **Sandbox environments**: Open Banking provides access to a third party sandbox environment that enables approved participants to build an understanding of Open Banking tools, standards and security requirements. Similar approaches could be adopted elsewhere.
- **Ring fencing data**: a legal or regulatory barrier segregating a portion of data from other data. This could maintain or increase competition in the market by, for example, preventing social media companies from combining consumers' social media data with their financial data. Further information is outlined in Box 3: Data Exchange Reciprocity.[63]

Available Resources

Decision-makers designing, implementing, and monitoring Smart Data schemes will require technical, legal, and financial resources to do so. Those governing schemes will need adequate resources to monitor and shape the scheme after its initial launch. Those working on a Smart Data scheme should look to draw lessons from other Smart Data schemes, both nationally and internationally where appropriate. Those developing Smart Data schemes should be conscious of the technical, legal, and financial burden that implementing a scheme might have on data holders and Authorised Third Parties. Considering the following questions could enable decision-makers to ensure that the Available Resources are sufficient in a Smart Data scheme:

Decision-maker:	**Questions for decision-makers to consider:**
Government Department(s)	• Is there an existing regulator or regulators that could oversee the introduction, and ongoing monitoring, of the scheme? ○ Does this regulator or regulators have sufficient resources to do so?
Regulator(s)	• Is there an existing organisation well-placed to act as the governing body of this scheme? Or is a new organisation

[62] Competition and Markets Authority, *Retail banking market investigation*, February 26, 2016.
[63] Amy Borrett, "How Australia is challenging the UK on open banking", *The New Statesman*, April 14, 2021.

	required? ○ How could this organisation be funded? (For example, the OBIE was funded by the UK's nine largest current account providers). ● Who should be responsible for maintaining and governing the scheme after its launch? ○ How could they be funded?
Scheme Governing Body	● Have you fully explored the potential lessons from other Smart Data schemes? (This includes those in the UK, internationally, those in development, and those already launched, and more general work on the subject, such as that produced by the SDWG[64]). ○ Is there an accepted international best practice relating to schemes in your sector? ● Have you fully examined how the nuances of your sector will impact the requirements of a scheme? ● Do you have sufficient in-house technical expertise to design the appropriate standards? ● Do you have sufficient legal expertise attached to your project? ● Will data holders/ Authorised Third Parties in your sector be well placed to participate in schemes?
Data Holders/ Authorised Third Parties	● Do you have sufficient technical expertise to fulfil the requirements of the scheme? ● Do you have sufficient legal expertise to fulfil the requirements of the scheme? ● Do you have sufficient financial expertise to fulfil the requirements of the scheme?
Cross-sectors considerations:	● Which regulator/regulators and governing body/bodies could be responsible for overseeing the introduction, and ongoing monitoring, of cross-sector data sharing? ○ Does this regulator/regulators or governing body/bodies have sufficient resources to do so? ○ Are additional resources required to facilitate coordination where multiple regulators are involved? ● Will organisations from different sectors be capable of handling data sharing with organisations from another sector? (Technology oriented organisations will likely be better placed to handle data sharing than others.)

Possible interventions
Those designing Smart Data schemes could explore existing toolkits, draw lessons from other existing Smart Data schemes, and engage with legal and technical support throughout the design and implementation phases. Aside from this, they could consider the following interventions:

- **Sandbox environments for development**: see above: Clear Scope, *Possible interventions*.
- **Funding routes for Authorised Third Parties**: Some Authorised Third Parties,

[64] https://www.gov.uk/government/groups/smart-data-working-group

particularly those providing goods and services to consumers in vulnerable circumstances, may require additional funding in the form of patient capital, challenge funds, etc. to allow these providers to compete against better-resourced businesses.

Effective Governance Mechanisms

Decision-makers should consider what effective governance and oversight mechanisms might be necessary to support their Smart Data schemes. This will likely include establishing clear lines of appropriate responsibility, with routes to redress if issues arise, outlining who is responsible for holding participating organisations accountable, and creating oversight mechanisms.

The decision-makers for a scheme will need to consider how to monitor and evaluate participation, and how to provide a clear process for redress where boundaries are crossed or a system is compromised. This includes making lines of appropriate responsibility clear to everyone involved with these schemes, particularly consumers. In addition, it will be important to incorporate diverse opinions and perspectives in the governance of Smart Data schemes. Considering the following questions could help decision-makers implement Effective Governance Mechanisms in Smart Data schemes:

Decision-maker:	**Questions for decision-makers to consider:**
Government Department(s)	● Does the scheme require more than one regulator to be involved in decision-making? ○ What arrangements might be necessary for coordination or cooperation?
Regulator(s)	● Have you established and communicated the overarching governing principles for your scheme? ● What governance and oversight mechanisms will be needed to govern the scheme? ○ For example, the CMA created the OBIE in 2016 to create the software standards and industry guidelines for Open Banking. In March 2022, the FCA, PSR, CMA and HM Treasury published a joint statement on the future of open banking, announcing the creation of a joint regulatory oversight committee.[65] ● Should Authorised Third Parties require accreditation to participate in the scheme? (See below: Box 5: Accreditation) ○ How should organisations become accredited? ○ What level of accreditation should be necessary? ○ Who should be responsible for accrediting relevant institutions? ○ How frequently should accreditations be reviewed? ○ What recourse should exist for violations by accredited parties? ○ What should the process be for removing accreditation?
Scheme Governing Body	● Do clear lines of appropriate responsibility exist? ○ How are these lines of responsibility communicated to consumers?

[65] https://www.fca.org.uk/firms/future-open-banking-joint-regulatory-oversight-committee

	● Are there clear liability mechanisms in place to ensure that inappropriate behaviour is open to redress? ○ Is it clear who is responsible for addressing consumer complaints? ○ Do consumers understand their routes to redress? ● How do you ensure that consumers provide consent to data sharing and processing? ○ How are consent mechanisms determined? ○ What information should be communicated to consumers when securing consent? ○ How frequently should consent/re-consent be required? ○ How do you ensure consumers are aware they can remove consent? ● Have you ensured there is enough representation of consumers and small/medium enterprises (SMEs) within the governing structure of the scheme?
Data Holders/Authorised Third Parties	● Do you have adequate internal governance structures to ensure you comply with the relevant legislation? ● Are your internal governance activities recorded end-to-end to ensure auditability? ● How do you communicate lines of liability and routes to redress to your consumers? ● How will you secure consent from consumers? ○ What information will you communicate to consumers when requesting consent? ○ Is it at least as easy to withdraw consent as it was to give consent? ○ How will you inform consumers that they can withdraw consent?
Cross-sector considerations:	● What organisation/organisations should be responsible for regulating cross-sector data sharing? ● What additional governance and oversight mechanisms will be necessary to govern the scheme? ● Would cross sector data sharing affect the lines of accountability and liability structures associated with data sharing? ● Would cross-sector data sharing affect the manner in which consumers consent to participation? ● If accreditation is deemed necessary, who should be responsible for accrediting participants in cross-sect0r schemes? ○ Should organisations be accredited by the authorities that govern each individual sector? ○ Should a central, single body manage accreditation across schemes/or for all schemes more generally?

Potential interventions

Decision-makers could suggest, mandate, or provide the following interventions that relate to lines of advice, accountability, and responsible data management:

- **Regulatory advice**: The Open Banking Standard provides guidance for

organisations that are implementing the Standard. They break this guidance into three sections: Mandatory (required in all cases), Conditional (required in some cases), and Optional.[66]

- **Internal lines of accountability**: Just as there should be clear lines of accountability between consumers, Authorised Third Parties and data holders, it should be clear who is responsible for what within organisations that handle data.
- **Data stewardship principles**: these are principles by which individuals and organisations act when handling data.

Box 5: Accreditation
The majority of existing Open Banking schemes require that organisations have licences to participate. Two exceptions are the US and Hong Kong. The Hong Kong Monetary Authority (HKMA) is currently rolling out its Open API Framework.[67] Whilst there is an expectation that retail banks will adopt this framework, it is not mandatory for them to do so. Alongside not being mandatory, there is no requirement that participating organisations receive accreditation. Accreditation in Hong Kong is different from Open Banking in the UK, where participants must enrol with the OBIE, which maintains a directory of whitelisted participants. Participants must be whitelisted prior to being allowed to provide Open Banking-enabled services. Stakeholders frequently mentioned accreditation mechanisms as necessary for building public trust in the Smart Data ecosystem. Some stakeholders argued that use of a kitemark could be a way of securing some consumers' trust in the schemes. They pointed to kitemarks such as the BSI Kitemark for product testing and the HTTPS lock icon as useful examples.

Meaningful Engagement

Effective communication with domain experts, affected stakeholders, and the general public will be important for fostering trust in Smart Data schemes. Stakeholders emphasised the importance of a multi-stakeholder approach, ensuring discursive forums between different groups that do not usually interact. Stakeholders we interviewed that were particularly involved in Open Banking highlighted that Open Banking could have had more active consumer representation in its development and implementation phases. Those developing Smart Data schemes should engage with a diverse set of stakeholders, particularly consumers with vulnerabilities and traditionally underrepresented groups. The results of this engagement will be useful in informing the design and implementation of these schemes. Likewise, continued engagement with consumers, after the implementation of the scheme, will aid in efforts to ensure a scheme is governed appropriately. Considering the following questions could help decision-makers implement Meaningful Engagement in these schemes:

Decision-maker:	**Questions for decision-makers to consider:**
Government Department(s)	● Have you communicated the underlying purpose of the Smart Data scheme to all potential parties in the scheme? ○ Have the potential *benefits* of the scheme or schemes been communicated?

[66] Open Banking, *Guidance when implementing the Standard*, accessed September 27, 2021.
[67] Hong Kong Monetary Authority, *Open Application Programming Interface (API) for the Banking Sector)*, accessed September 27, 2021.

	○ Have the potential *risks* of the scheme or schemes been communicated?
Regulator(s)	● Have stakeholder groups (including: consumers, consumer groups, incumbent data holders, and Authorised Third Parties) been engaged at all stages, from design through to implementation and evaluation? ● Is it easy for consumers to find out: ○ The rights they have in relation to their data? ○ The liability structures in the Smart Data scheme? ○ The available routes to redress if things go wrong? ○ How they can revoke consent for their data to be shared?
Scheme Governing Body	● Have stakeholder groups been engaged at all stages from design through implementation? (see above: Regulator(s)) ● Have you incorporated the results of this engagement into the development and implementation of the scheme? ● Which consumers are considered vulnerable within your sector? ○ How have you engaged with, and incorporated, the needs of consumers in vulnerable circumstances into your scheme? ● Have the roles and responsibilities of your body been clearly communicated to consumers?
Data Holders/Authorised Third Parties	● Have you adequately communicated the following to consumers: ○ The benefits offered to them by the scheme? ○ The risks of the scheme? ○ What elements of their data will be used? ○ What are their rights in relation to their data? ○ How their data will be used? ○ Who will use their data? ○ How long their data will be used/stored? ○ How they can opt out of the scheme? ○ Their routes to redress?
Cross-sector considerations:	● How could you engage consumers when determining if cross-sector data sharing is desirable? ● How will allowing cross-sector data sharing affect which consumers are considered vulnerable? ● Do consumers understand the following, particularly given the complexities that cross-sector data sharing will add: ○ The benefits of using elements of their data from multiple sectors? ○ The risks of using elements of their data from multiple sectors? ○ What elements of their data will be used, and where these elements come from? ○ What are their rights in relation to their data? ○ How their data will be used? ○ Who will use their data? ○ How long will their data be used/stored?

	○ How can they opt out of their data being used? ○ Their routes to redress?

Possible interventions
The following convening- and consultation-based interventions could help ensure that meaningful engagement is undertaken whilst designing and implementing a Smart Data scheme:

- **Public attitudes polling**: For example, the CDEI undertook some public attitudes polling to understand what might prevent consumers from participating in Smart Data schemes[68];
- **Expert convening**: this could vary from individual interviews with experts to scenario planning exercises;
- **Public consultation**: this could include actively seeking out input from the public on specific questions whilst developing the scheme;
- **Engagement with special interest or civil society groups**: organisations that represent groups that have been underrepresented elsewhere in engagement, particularly groups with vulnerabilities, should be sought for their expertise and perspective.

Implementing ethical and trustworthy Smart Data schemes in practice

Table 4: Example decision points for Smart Data schemes provides the CDEI's initial thinking on implementing Smart Data schemes in a trustworthy and ethical manner, leading on from the features, conditions, and interventions identified in Sections 2 and 3. It highlights that ethical and trust-related challenges could be addressed at all relevant decision points and by all relevant decision-makers when establishing these schemes.

The table provides an idea of how a scheme could be developed, not necessarily how one *must* be developed. For example, whilst the table distinguishes between regulators and governing bodies, it is not a given that a regulator and governing body will be different organisations within a scheme (or schemes). There may also be sectors where there is no specific regulator, or several regulators, where multiple organisations will need to coordinate outcomes. The unique context surrounding each scheme matters and will have a significant impact on its development and implementation. The table imagines a scheme being implemented in one specific sector, where a government department tasks a regulator, who in turn tasks a governing body, with developing the scheme. This work is high level, and a more full version of thinking is given in the Phase 3 of the CDEI's work (Smart Data Implementation Guide).

[68] The CDEI, Examining public attitudes towards Smart Data schemes, June 2022.

Table 4: Example decision points for Smart Data schemes

	1. Set up and planning	2. Develop and test	3. Voluntary onboarding and testing	4. Onboarding and services go live	5. Transition to business as usual
Government Department(s)	- Prove high level case for the scheme. - Determine regulator/group responsible for scheme				- Consider if scheme could be linked to other Smart Data schemes
Regulator(s) (e.g. CMA, FCA)	- Work out core features and driving needs (e.g. competition) behind scheme - Design top key features/remedies of scheme - Decide the governing body responsible for designing standards and implementing the scheme - Determine initial funding for Governing Body - Consult relevant stakeholders - Establishing governance framework	- Input into Governing Body's standards	- Input into Governing Body's standards - Test safety and security - Establish regulatory framework	- Regulate participants of scheme	- Regulate participants of scheme - Test safety and security - Continue communication, and listening to, consumers (particularly in regards to liability structures) - Maintain clear communication routes for consumers and consumer groups, continuously taking their input onboard - Respond to and act upon consumer complaints
Governing Body/Bodies (e.g., OBIE)		- Develop standards and guidance around necessary data architecture - Incorporate trustworthiness and data rights into design - Decide rollout phases - Decide security standards - Determine accreditation requirements and process - Design liability framework - Consult relevant stakeholders - Create test environment	- Adjust standards and guidance as necessary - Decide security standards - Determine accreditation requirements and process - Design liability framework - Consult relevant stakeholders	- Rollout scheme in phases - Maintain standards - Maintain guidance for participants - Maintain accreditation directors - Test safety and security - Communicate liability structures to consumers - Develop any centrally run infrastructure	- Maintain standards - Maintain guidance for participants - Maintain accreditation directors - Test safety and security - Continue communication, and listening to, consumers (particularly in regards to liability structures) - Maintain clear communication routes for consumers and consumer groups, continuously taking their input onboard. - Respond to and act upon complaints
Data Holders /Authorised Third Parties		- Engage with regulator and Governing Body - Participate in beta testing	- Engage with Governing Body - Develop any desired additional standards not mandated by the Governing Body deemed to be helpful - Participate in pilot phases - Develop propositions based on scheme	- Fulfil participation/accreditation criteria - Construct necessary architecture - Ensure safety and security of architecture - Communicate value of scheme to consumers - Seek consent from consumers - Communicate liability structures to consumers	- Fulfil participation /accreditation criteria - Ensure safety and security of architecture - Communicate value of scheme to consumers - Ensure consent from consumers - Continue communication with consumers (particularly in regards to liability structures) - Develop new offerings/services
Consumers/ Consumer Groups	- Engage with Government Department and Regulator	- Engage with Regulator and Governing Body	- Engage with Governing Body	- Sign up to schemes - Provide ongoing consent	- Sign up to schemes - Provide ongoing consent - Continue providing feedback to all of the above parties

Section 4. Next phases and areas of future work

The CDEI has been working with 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, desk based research, and studying of international approaches to Smart Data.
- 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 paper represents the findings of the first phase of work the CDEI has undertaken with BEIS, which has involved identifying the features, conditions, and interventions that could underpin ethical and trustworthy Smart Data schemes. This paper has been developed through engagement with stakeholders and desk research of existing Smart Data schemes nationally and internationally. It was used to inform the next two phases of this project.

Phase 2 - Envisioning the future of Smart Data schemes in 2028

In Phase 2 of this work, the CDEI delivered a report based on a Scenario Planning and Visioning workshop undertaken with a diverse group of stakeholders, including relevant government departments, regulators, data holders, Authorised Third Parties, and consumer representatives. Scenario Planning is a futures methodology that is used to discuss how present-day uncertainties could develop into multiple different futures, helping to build consensus around what positive and negative outcomes could look like for Smart Data Schemes. Doing so helped us to identify what roles different stakeholders could and should play in these schemes and to ensure that any decisions taken today are robust to potential future changes. Practically, this could help in identifying steps that should be taken to drive positive outcomes, such as stimulating innovation and driving fairer outcomes for consumers.

The results of this workshop can be found in the Phase 2 “Scenarios Report: the Future of Smart Data 2028” report.

Phase 3 - Smart Data Implementation Guide

In Phase 3 of this work, the CDEI and BEIS engaged with a diverse sample of consumers to develop a deeper understanding of public perceptions of Smart Data through workshops. The workshops focused on assessing differences in consumer sentiment within and across sectors. For example, do consumers feel more concerned about sharing their financial data versus their energy data? What accounts for why these concerns vary? Further to this, the workshops also focused on understanding how consumers consider trade-offs between features of Smart Data schemes. For example, how do consumers balance the benefits of Smart Data-enabled services against their privacy concerns? Finally, consumers were asked to assess the scenarios developed by government, industry, and third sector stakeholders to test the validity of underlying assumptions made by these stakeholders.

This work was used to inform the implementation guide developed in Phase 3 of the CDEI’s work.

Appendix: List of Interviewees

The CDEI and BEIS are incredibly grateful to all that gave their time to be interviewed as part of this work. A number of these interviews were carried out under the condition of anonymity.

Some of the organisations the CDEI spoke to were:

- Broadband UK
- BT
- The Competition and Markets Authority
- The Department for Digital, Culture, Media & Sport
- Expedia
- The Finance Innovation Lab
- The Financial Conduct Authority
- The Financial Inclusion Centre
- HSBC
- Icebreaker One
- Lawtech UK
- Money & Mental Health
- NatWest
- Ofcom
- The Open Banking Implementation Entity
- The Open Data Institute
- Plaid
- Swoop Funding
- TrueLayer
- The University of Nottingham
- Which?

Bibliography/Further Resources

Access Now, *Data Minimization: Key to Protecting Privacy and Reducing Harm*, accessed September 27, 2021.

Ada Lovelace Institute, *Exploring principles for data stewardship*, September 18, 2020.

Amy Borrett, "How Australia is challenging the UK on open banking", *The New Statesman,* April 14, 2021.

Ariadne Plaitakis and Stefan Staschen, *Working Paper: Open Banking: How to design for financial inclusion,* 2020;

The Centre for Data Ethics and Innovation, *Addressing trust in public sector data use*, July 2020.

The Centre for Data Ethics and Innovation, *PETs Adoption Guide*, 2021.

The Centre for Data Ethics and Innovation and Department for Business, Energy, and Industrial Strategy, *Smart Data: Open Finance and Open Communications: Public Attitudes Survey,* 2022.

Centre for Data Ethics and Innovation, *Smart Data schemes: Landscape review discussion note*, July 2021.

Competition and Markets Authority, *Consultation launched on the future governance of open banking*, March 5, 2021.

Competition and Markets Authority, *Retail banking market investigation*, February 26, 2016.

Department for Digital, Media, Culture, and Sport, *National Data Strategy*, December 2020.

Department of the Treasury, the Australian Government, *Consumer Data Right*, May 9, 2018.

Department of the Treasury, the Australian Government, *Consumer Data Right Overview*, September 2019.

Department of the Treasury, the Australian Government, *Open Banking Guidelines for Open Data Participants*, July 2018.

Faith Reynolds and Mark Chidley, *Consumer priorities for open banking*, 2019.

Financial Consumer Agency of Canada, *Open Banking*, 2021.

Financial Conduct Authority, *Finalised Guidance: FG21/1 Guidance for firms on the fair treatment of vulnerable customers*, February 2021.

Han-Wei Liu, "Two decades of laws and practice around screen scraping in the common law world and its Open Banking watershed moment", *Washington International Law Journal* Vol 30. No. 1, 2020.

Hong Kong Monetary Authority, *Open Application Programming Interface (API) for the Banking Sector)* accessed September 27, 2021.

ICO, *Right to data portability, accessed September 27, 2021.*

IETF, *Terminology for Talking about Privacy by Data Minimization*, January 6, 2011.

Matt Warman MP and DCMS, *Next steps in plans to govern use of digital identities revealed*, August 2, 2021.

Open Banking, *Guidance when implementing the Standard*, accessed September 27, 2021.

Open Data Institute and Fingleton Associates, *Data sharing and Open Data for banks: A report for HM Treasury and Cabinet Office*, September 2014.

Sara Davies, Andrea Finney, and Yvette Hartfree, *Paying to be poor: Uncovering the scale and nature of the poverty premium*, November 2016.

Smart Data Working Group, *Smart Data Working Group: Spring 2021 report*, June 2021.

Smart Data Working Group, *Next Steps for Smart Data*, March 2021;

The Association of Banks in Singapore, *Singapore Financial Data Exchange (SGFinDex)*, accessed September 27, 2021.

University of Bristol, *The poverty premium: When low-income households pay more for essential goods and services*, November 2016.

[inside of the back cover – for printed publications, leave this page blank]