Smart Data Research UK

UK Research and Innovation

Consultation response:

Data intermediaries

May 2025

Introduction

This response addresses the UK government's call for evidence on data intermediaries, which recognises the transformative potential of data in driving innovation and economic growth. As the consultation notes, data is integral to people's lives and fuels most modern business, with data-specialised businesses increasing their share of GDP from 6.5% to 7.4% between 2021 and 2023 alone.

Data intermediaries are organisations that facilitate access to and exchange of data, acting on behalf of or for the benefit of data subjects when dealing with personal data. Unlike data brokers, they rely on the agreement of individuals and act in their interest. They can play a crucial role in enabling data subjects to exercise their rights, particularly the right to data portability, by helping individuals move their data from one data controller to another.

However, significant opportunities remain untapped in the broader data economy. Only 21% of businesses that handle digitised data analyse it to generate new insights, and just 14% share data outside their organisation. Data intermediaries represent a promising pathway to unlock this potential while ensuring individuals retain control over their personal data.

Smart Data Research UK (SDR UK) welcomes this consultation as timely and important. As a new UKRI investment in data infrastructure, we are funding work that strengthens R&D resulting from data intermediary services. The example that we highlight particularly in our response is the Smart Data Donation Service and its work with a Trusted Research Environment. Our response draws on practical insights from building research data services and the barriers we have encountered in enabling individuals to exercise their data rights for research purposes.

The consultation's focus on trusted and secure data use aligns closely with SDR UK's mission to unlock the potential of digital data for research while maintaining the highest standards of data protection and public trust. Our response contributes evidence on how data intermediaries can operate effectively within the current regulatory framework and where additional clarity is needed to realise their full potential.

Data Intermediaries: Call for Evidence

All consultation documents are available here: Data intermediaries - GOV.UK

Section A: Exercise of data subject rights

Q1. Can you provide examples of where data subject rights are currently exercised by third parties on the instruction of, or in the interest of, the data subject?

Smart Data Research UK (SDR UK) is a new UKRI investment in data infrastructure aiming to unlock the potential of new forms of digital data for research. Such data holds immense potential to generate insights, solve social and environmental challenges, and drive economic growth.

We strongly agree that data intermediaries have the potential to empower individuals to exercise their data subject rights, and providing legal clarity about the legitimate role of intermediaries could help improve incentives, and reduce frictions, for citizens to exercise their rights. Our response to this call for evidence also seeks to represent some of the barriers to address towards achieving that approach.

To date, the majority of research undertaken with smart data has not directly invoked data subject rights (e.g. UK GDPR Article 15, Article 20). There have been individual studies where participants have been recruited and specifically asked to access their digital device data [1]. However, developing high-quality data assets for wide-scale research use in this context is challenging. This is – in part – because significant

effort and expense is required to develop and maintain the infrastructure necessary to support such the large-scale acquisition, cleaning, and enrichment of such data. Efforts are hampered by legal and regulatory uncertainty – see below – as well as by an uneven corporate landscape. Significant effort (and bespoke legal advice) is often required (for data subjects, as well as for research institutes) to access data and to use it in the appropriate ways- either from companies or from data subjects making use of their services.

The opportunity presented by data portability rights, as well as the constraints to researchers' ability to benefit from these rights, are drivers behind SDR UK's investment in a data service focusing specifically on 'data donation'. SDR UK's Smart Data Donation Service, led by researchers at the University of York, aims to empower citizens to obtain and share their own digital data safely with researchers. The service will integrate data donation into pre-existing national data assets (e.g. Understanding Society, Born in Bradford); and independently create novel data assets. Together, these strategies aim to enable wider use of donated data to address questions regarding online behaviour and its relationship to wellbeing.

This initiative aims to facilitate both (a) best-in-class methodological research to assure the quality of data obtained via data subject rights; and (b) domain-specific research that utilizes digital record data (e.g. regarding social media use and gaming) to drive evidence generation across a variety of domains. It potentially offers powerful new insights into how people behave, trends in the population, and can also inform policy making.

Footnote:

[1] E.g. in a study which looked into retail shopping data, the Cancer Loyalty Card Study succeeded in recruiting a cohort to look into symptoms and time-to-diagnosis of ovarian cancer (https://www.clocsproject.org.uk/).

And, e.g., to obtain data from video gaming platforms, the Oxford Internet Institute recruits participants to share their data for research projects focusing on video game play and mental health, by working in partnership with the video gaming industry (https://www.oii.ox.ac.uk/research/projects/understanding-video-game-play-and-mental-health/).

Q2. What barriers do individuals, businesses, or other organisations face in the uptake of the right to data portability or other data subject rights?

Legal and regulatory uncertainty
For research data access that is specifically to follow the exercise of data portability rights, it's not particularly defined how a third party (e.g. SDR UK's Smart Data Donation Service) will demonstrate that they exercise data subject rights on behalf of an individual.

There is a general understanding that the UK GDPR does not prevent third-party organisations from invoking data rights on the behalf of an individual. However, in practice, robust and standardised practices associated with frictionless third-party enactment of such rights by data intermediaries (e.g. data services) are currently under-defined.

Official guidance suggests that it is necessary for organisations to assure themselves that an intermediary is acting on the behalf of an individual before responding to a data rights request, and gives broad categories for evidence that may be used to substantiate this, subject to verification (e.g. power of attorney). However, there is no explicit and concrete specification which could lead to an unambiguous process. Endorsement of a standardised mechanism for official authentication and authorisation with reference to a specific data rights request, would resolve this barrier at scale. The lack of such standardisation represents a key limit to data rights uptake in this domain.

And, in alternative routes, when researchers request data from companies: when data is sought, it is

difficult at first to establish whether data will flow because drivers towards data sharing are not particularly enforced. For example, among researchers seeking to make statistical use of data from the UK retail sector, there's often a reported lack of capacity or incentive to share data to the standard that we'd expect.

Uneven corporate landscape
Implementation of the right to data portability is highly uneven across the corporate landscape. Whilst some companies (e.g. Google, Tesco) provide low-friction access to data downloads and API-based data transfers, other competitors provide minimal guidance to individuals wishing to enact their data right (e.g. an email address to which individuals can write). Some providers will action requests in a matter of minutes via an automated service; others require up to a month. This limits the ability of individuals to benefit from their data rights, and also acts as a ceiling for researcher access to data.

Unsecured Data Fragmentation
The results of data portability requests and data Subject Access Requests frequently consist of the kinds of granular and rich time-series data that are required to answer the most pressing questions in multiple fields. However, the richness of these data mean that they are also frequently capable of identifying an individual unless subjected to substantial data processing. Researchers must choose whether to reduce the complexity (and hence use/reuse value) of these data to render them deidentified via such processing; or deal with the risks associated with holding such potentially hazardous data. Few individual labs are independently capable of such risk management.

Therefore, widespread utilisation of data subject rights as a data acquisition strategy across the research ecosystem may lead to the creation of myriad under-secured datasets, potentially creating novel risks to individuals as a consequence of enacting their data rights. At SDR UK, we aim to mitigate these risks via data centralisation within well-secured environments (e.g. the Smart Data Donation Service and its associated TRE).

Barriers to data enrichment
When data is accessed via a data portability request, resulting outputs vary in terms of their ability to be enriched. For example, some corporations provide outputs in the form of resource identifiers or locators which cannot be enriched for research purposes without breaking that platform's terms of service. This may technically be GDPR compliant, but limits both the individual and the wider research community from benefitting from this right.

Consent mechanisms
Research teams obtaining data on the GDPR legal basis of consent, need to maintain the right level of communication with participants and enable their consent and their right to data portability to continue. Dynamic consent mechanisms are needed because the individual's position about what their data is provided for can change.

For example, the Energy Systems Catapult (a partner in SDR UK's Smart Energy Data Service (SENSE)) obtains smart meter data by way of consent, and returns for re-consent to the research use, making clear that participants can withdraw from future studies.

Anonymisation
In Trusted Research Environments that are focused on dealing with public sector data, it is not only the security of the data but its level of anonymisation which is an important pre-requisite for the public to trust expansion in research use. We find that the same is true for research making use of sensitive data that is drawn from the private sector [2].

TREs that are operating on this basis do not assume the roles and obligations of a data controller and/or processor, rather they assume the role of receiving de-identified data, and vetting responsible research, and responsible research practices, in the use of the data. Attributes of sensitive data can be picked up for analysis within the TRE, and the data which is accessed for a statistical research purpose is either pseudonymised, or completely anonymous.

For example: Smart Data Foundry (a founding partner in SDR UK's new Financial Information Data Service, FINDS), receives financial transaction data only after it has been effectively anonymised and separated from the original records.

Public trust and transparency
Uses of data can be controversial to the public as well as to firms whose activities are reflected (even if anonymised) in shared data.
Mitigating public concerns requires clarity about, e.g., acceptable incentives to take part in research, and ethical and legal guardrails. To satisfy the premise for data sharing, it must be clear that intermediaries are operating in the interests of the data subject, and that significant vulnerabilities for participants are not opened up, either for the individuals or the wider public. Data subject rights will need to be balanced with rights of other individuals in the research.

For example, biomedical data services operating in the UK, are managing data from large cohorts of donors, and approve a growing base of users from many different types of organisation (e.g., for-profit as well as non-profit). Governance approaches involve participant representatives to guide how the data and its use interacts with public concerns. Ethical guardrails are communicated to all participants, whose support is integral to all future activities. [3]

Establishing governance, ethical guidelines and public engagement is an upfront cost, and beneficial to clarify. SDR UK is an active partner in the Public Engagement in Data Research Initiative, and we dedicate portions of our strategic hub resource towards public engagement as well as planning ethical and legal support, coordinated with data services in the newly funded programme.

Footnotes:

[2] Smart Data Research UK (2025) Public Dialogue 2025. May 2025 - https://www.sdruk.ukri.org/wp-content/uploads/2025/05/SDR-UK-Public-Dialogue-report-2025-1.pdf

[3] E.g.: Our Future Health [website] 'How we control access to data' https://ourfuturehealth.org.uk/protecting-your-data/how-we-control-access-to-data/

And, e.g. Genomics England [website] 'The Participant Panel' https://www.genomicsengland.co.uk/patients-participants/participant-panel

Q3. Aside from personal data protection laws, how do other areas of law interact with the operation of data intermediaries?

The specific needs to foster data standards when sharing consumer data are recognised in industry regulatory frameworks, particularly so in Open Banking.

In connection with approaches like open banking that focus particularly on businesses and their need for data standards, research organisations are participating (in some instances) as recipients from the third parties.

An established partnership between a secure research data service and a social enterprise, Salad Money, provides one example of how research use can take place further down the line from Open Banking, after data has been anonymised. Salad Money launched in 2019 to give NHS and public sector workers with poor or thin credit scores an alternative to high-cost lending. Salad Money receives around 50,000 applications for credit every month and analyses around 1,500 bank transactions for each applicant to make its lending decisions. Their Data Licensing Agreement with the Consumer Data Research Centre (based at Liverpool University, Oxford's Saïd Business School, and UCL) agrees terms for researchers to access anonymised Open Banking data about the use of consumer credit by workers, for analysis and for research, within a secure data service. This has enabled researchers throughout the UK to evaluate issues such as, financial vulnerability and the impact of loans. [4]

In a different industry sector, regarding energy systems data, UK regulator Ofgem is developing practices that support data sharing in a complex and fragmented industry, through data licence conditions.

Where there are industry models for standardising and connecting data, then the receipt of readily-usable, operator data, in addition to the development of consumer/public surveys, is of great help to characterise individual, population-level, and geographic data.

In connection with the regulation of data sharing, cases for research use could more particularly be recognised and guidance clarified. There are some examples of this in other areas of law, e.g.:

- to access data from Very Large Online Platforms, the EU's Digital Services Act is just coming into force and stipulates a range of requirements for researcher access through a delegated act.
- In domestic (UK) legislation, for researchers to access non-health data from central government departments, the Digital Economy Act 2017 facilitates the linking and sharing of de-identified public sector data to support research and statistical work, underpinned by the UK Statistics Authority's Research Code of Practice and Accreditation Criteria for the public good.

Footnote:

[4] 'Unique collaboration will unlock insights into key-workers' financial situations and behaviour' [Salad Money website], About > News, 9th August 2023
https://www.saladmoney.co.uk/about/news/unique-collaboration-will-unlock-insights-into-key-workers-financial-situations-and-behaviour

Section B: Data intermediaries

Q4. Does the taxonomy above fully reflect the range of models of data intermediaries in the UK or elsewhere?

Secure data settings for research (Trusted Research Environments, etc) can stand apart from what is strictly termed personal data, depending upon what their relationship is to the individual's record and how they receive the data.

Q5. Is the current law around the operations of data intermediaries sufficiently clear? What changes and/or additional guidance would be required to provide clarity to data intermediaries? Does this differ based on operating model?

There is more than one way to comply with GDPR in regard to data sharing for research, and differences in experience and interpretation can cause confusion to researchers and prospective data suppliers.

Provisions that are made for scientific research in accordance with UK GDPR are not well known, and data-controlling businesses and intermediaries are, presently, cautious about the potential impact of ICO rulings (even when those rulings are geared towards regulating the use of data for purposes other than research – such as direct marketing and political campaigns). This can present as reluctance to engage in long term data licencing or intermediation agreements which could provide more favourable terms for research uses to develop.

In addition to GDPR's complex application, data sharing for research is also affected by other laws:
Competition law.

This plays a role in enforcing data portability principles and affects the extent to which intermediaries' activities can be contested by companies.

Intellectual property and copyright
Data subjects' access to data and their data portability right needs not to infringe copyright or IP rights held by the companies involved.

Consumer rights frameworks, and legal provisions for researcher access to consumer/user data.

And, in regard to sharing of digital health data or patient records:

- Common Law Duty of Confidentiality and the COPI Regulations
- NHS Act 2007

And so, there is scope for legal interpretation to clarify that intermediaries can make data available for research purposes under terms guiding the secure research operating models that fully accord with UK GDPR. A general model for federating Trusted Research Environments to enable data linkage would, additionally, be helpful, and would bring about greater value from the data.

[Response ends].

Smart Data Research UK

UK Research and Innovation

Stay in touch

Email us
info@sdruk.ukri.org

Find us
sdruk.ukri.org

Bluesky
sdruk.bsky.social

LinkedIn
Smart Data Research UK