Data Sharing and Linkage for the Public Good: Follow-Up Report

July 2024

Office for Statistics Regulation

We provide independent regulation of all official statistics produced in the UK. Statistics are an essential public asset. We aim to enhance public confidence in the trustworthiness, quality and value of statistics produced by government.

We do this by setting the standards they must meet in the [Code of Practice for Statistics.](#) We ensure that producers of government statistics uphold these standards by conducting assessments against the Code. Those found to comply with the standards of trustworthiness, quality and value in the Code are given accredited official statistics status (called National Statistics in the Statistics and Registration Service Act 2007). We also report publicly on system-wide issues and on the way statistics are being used, celebrating when the standards are upheld and challenging publicly when they are not.

Acknowledgements

We would like to acknowledge the significant input of all the individuals and organisations that contributed to this review as we gathered information and tested the ideas presented here. Organisations and teams that contributed are listed in Annex A.

Contents

Summary

In 2023, the Office for Statistics Regulation (OSR) reviewed the data sharing and linkage landscape across government and made 16 recommendations to enable greater data sharing and linkage for research and statistics for the public good. This follow-up report assesses the progress that has been made towards those recommendations. It is relevant for all who want to see good-quality and secure data at the forefront of government decision-making and research.

Despite welcome pockets of innovation, there continues to be a failure to deliver on data sharing and linkage across government, alongside many persisting barriers to progress. Linking datasets for research, statistics and evaluation – both across government and among external researchers – is not yet the norm in the UK statistical system. To make this a reality, stronger commitments to prioritise data sharing and linkage are required. Such commitments further need to be endorsed and sustainably resourced by senior political and Civil Service leadership.

The formation of a new government in the UK offers the opportunity for renewed energy behind a cross-government, long-term strategy to ensure effective data sharing and linkage for the public good. This report outlines some of the actionable steps that can be taken towards that ambition and highlights examples of good practice in data sharing, access and linkage over the last year.

Definitions

Data Sharing and Access

The concept of data sharing is relatively straightforward: it is when data normally created in one business area or organisation moves to another. Data access is a form of data sharing where organisations contribute data to IT platforms, such as databases or modern cloud repositories. These platforms can then enable access to multiple others, removing the need for the source organisation to repeatedly share data themselves. These could be single datasets or data assets that have been prepared by linking multiple datasets. This is how a lot of data sharing is done across government. Data sharing and data access often rely on organisations having a common purpose and arrangements, such as an agreement to share data. In this report, we use 'data sharing and linkage' to mean both data sharing and data access, except where we specifically refer to one or the other.

Data Linkage

Data linkage involves bringing together separate datasets by identifying and matching the same entity in each using unique identifiers and then combining different datasets into a single dataset. Data that are shared between organisations are often shared with the intention of linking them to further datasets to enhance or improve the data. These new datasets can lead to new insights that weren't previously possible. In good data linkage examples, documentation is created and maintained to explain any relevant methods and data sources used in the creation of the linked dataset. This documentation can also help users understand the linked dataset and use it effectively.

Government

Throughout our report, we use the term 'government' to refer to the UK Government and the devolved administrations of Wales, Scotland and Northern Ireland. Where we are referring to a specific administration, this is made clear.

Office for Statistics Regulation

Overview

In July 2023, the Office for Statistics Regulation (OSR) published a [report on data sharing and linkage for the public good](#). The report reviewed the data sharing and linkage landscape across government and identified steps needed to move towards a brighter future for research and statistics, where data sharing and linkage is a priority across government and barriers have been successfully tackled. In the report, we made 16 recommendations which, if implemented, will enable greater data sharing and linkage for research and statistics for the public good.

Since the previous report was published, OSR has continued to engage with key stakeholders across government and beyond, aiming to generate and maintain momentum on our recommendations. Our report has had positive impacts on several projects relating to data sharing and linkage, as well as the culture in government around data use. These impacts include influencing the strategic approach taken by the Department for Science, Innovation and Technology (DSIT) to reviewing cross-government data sharing policy; developments in the [Data Marketplace](#) led by the Central Digital and Data Office (CDDO); the implementation of Wave 2 of the [Public Engagement in Data Research Initiative](#) (PEDRI); and technical innovation by the ONS Data Science Campus in developing new privacy-enhancing technologies (PETs). These initiatives reflect wider progress in several areas of data sharing and linkage, such as in public engagement activity and the support resources available to researchers and data users. Much of this positive activity has been led by National Institutes and programmes such as [ADR UK](#), [DARE UK](#) and [HDR UK](#), which help to bridge the gap between data owners and analysts in government and beyond. Nonetheless, there is still significant progress to be made in overcoming many of the remaining barriers to data sharing and linkage.

In our previous report, we committed to providing an update on progress against our recommendations within one year of publication. To assess progress, we have spoken to many of the same stakeholders, and some new ones, asking them for updates on work relevant to the recommendations. This report summarises our findings and our refreshed recommendations to government based on these findings. While some recommendations remain unchanged, others have been updated to reflect developments over the last year and to make them more useful to data owners and users.

As the regulatory arm of the UK Statistics Authority (UKSA), our interest in data sharing and linkage – and the focus of this report and its recommendations – primarily relates to research and statistics. We recognise that effective data sharing and linkage happens between some government departments to enable operational delivery of public services, and that there are indeed pockets of effective data sharing and linkage ongoing between departments for use in statistics. However, there is still much to be done across government to reach a place where sharing and linking datasets for research, statistics and evaluation – both across government and for external researchers – is the norm across the UK statistical system, rather than the exception.

Key findings

There continues to be a failure to deliver on data sharing and linkage across government. Leadership is needed to ensure data sharing and linkage between departments and with external researchers is prioritised. We are disappointed that progress on data sharing and linkage has been so slow across many areas highlighted in our previous report. The benefits of effective data sharing and linkage for the public good are clear. Sharing and linking data provides greater insights into society and the economy, enabling policy development and stimulating innovation, and is vital for socio-economic development. Without leadership across government, these benefits will not be realised and the UK will lag behind other comparable countries.

In March 2024, an independent review of the UKSA led by Professor Denise Lievesley CBE was published. We strongly endorse the recognition that "an effective system of responsible data sharing is critically important" to realise the potential of data held by government. The independent review also recognised the significant systemic, often cultural, barriers limiting data sharing across government. It recommended that the highest levels of government tackle these barriers and ensure data sharing for research and statistics purposes is prioritised. The cultural barriers were also highlighted in the May 2024 report Transforming the UK's Evidence Base by the Public Administration and Constitutional Affairs Committee (PACAC). For sharing and linking datasets to be the norm across the UK statistical system, rather than the exception, there is still much to be done across government. To move towards this aim, we need to see leadership from central government and the statistical system, concerted efforts from the wider data community and government working in partnership with external researchers and organisations.

There are promising signs that some parts of government are willing to take leadership around data sharing and linkage, but stronger commitments are needed. The Five Safes Framework underpins the government's approach to enabling researchers to access data for research and statistical purposes. Thus, it is vital to ensure that the framework continues to effectively enable data sharing and linkage in a secure way. The UK Statistics Authority (UKSA) and Research Accreditation Panel (RAP) have committed to engaging with our recommendation to review the Five Safes Framework. Progress here includes discussing the potential to broaden use cases for data by considering a change in focus from 'safe projects' to 'safe programmes'. The current strategic role of the Department for Science, Innovation and Technology (DSIT) in developing and cohering policy towards data access for the benefit of all of society and across the economy, has the potential to drive systemic change across government. In the government's formal response to the Lievesley independent review of the UKSA, the Cabinet Office has committed to push forward data sharing in government.

As a new government forms, we also recognise the role of politicians in enabling data sharing and linkage. Politicians should set the expectation that departments must work collaboratively on the issues facing society and the economy. The Cabinet Office should work with the new government to ensure that effective data sharing and linkage forms part of the new government's long-term strategy to tackle issues facing the UK. We look forward to the Cabinet Office's more detailed response to the Lievesley review later this year.

Senior leaders and all in government need more awareness of and exposure to data issues, but data literacy is not being prioritised. We previously pointed to the Data Masterclass – delivered by the ONS Data Science Campus in partnership with the 10 Downing Street Data Science Team – as a way for senior leaders to improve their data literacy. We suggested that this course be expanded to cover more topics, including the benefits of data sharing and linkage. So, it is disappointing that the Data Masterclass is currently on hold, with new cohort sign-up and course development paused. The importance of fostering a stronger data-driven culture in government was reflected in OSR's recent report, Analytical leadership: Achieving better outcomes for citizens.

The conversation around data sharing and linkage continues to focus on the risks – from the individual privacy risks to the reputational risks to data owners and government departments. Yet effective data sharing and linkage across government for research and statistics offers considerable benefits for the public good. It remains important to acknowledge the potential risks associated with sharing data and linked datasets, and to ensure both security and privacy. However, we would also like to see an emphasis on potential benefits for the public good. A UK-wide public dialogue report, produced by OSR and ADR UK, showed that the public support the use of government data in research and statistics, as long as they're informed of their use and what safeguarding is in place. Government should start by thinking about what it wants to achieve and focus on the potential positive outcomes from effective data sharing and linkage. Leadership, including political leadership, is needed to promote these benefits, and we see gaps across government in leadership on data sharing and linkage. This report also highlights how social licence and retaining public confidence remain fundamental for progress. More can be done to work with members of the public and raise public awareness of data as an asset which can benefit government as a whole.

There are pockets of innovative and ambitious work happening in government around data sharing and linkage and highly successful data sharing and linkage projects. The cross-government Better Outcomes through Shared Data (BOLD) programme continues to demonstrate how system-level developments can directly improve outcomes for the public. HMRC and ONS are expanding a long-standing data sharing arrangement for research and statistical purposes. In Northern Ireland, NISRA (Northern Ireland Statistics and Research Agency) has supported the growing interest in cross-departmental data insights, such as with the creation of the Educational Outcomes Linkage initiative. Meanwhile, new supporting materials have been introduced or progressed. The CDDO is developing a data marketplace to improve the discoverability of data within government and provides a central place for those within government to find out about what data are held and how data can be accessed. This departmental work is complemented, and often enabled, by the support of organisations such as ADR UK, which has launched a comprehensive Learning Hub bringing together information on skills and resources for those using administrative data and data linkage, as well as a searchable public metadata catalogue for datasets held across the ADR UK partnership. As for technical developments, there has been significant progress in the development and trial application of privacy-enhancing technologies (PETs) over the last year. DARE UK Phase 1 driver programmes are developing standards and new technical solutions to meet the needs of TRE users in relation to areas such as data analysis and

Office for Statistics Regulation

disclosure control. All these initiatives are positive and have the potential to support data sharing and linkage by government.

Processes for data sharing and linkage continue to create barriers to effective and efficient data sharing. Through discussions with stakeholders for this follow up, we have been made aware of additional process challenges. These include delays to output approvals by trusted research environments (TREs), barriers created by the requirement to publish all analyses enabled by the DEA (2017) Research Power and lengthy legal review of complex and non-standard data sharing agreements. The time taken to access data remains a significant barrier in some cases. Researchers using one secure research environment have reported significant increases in wait times for data access and output approval, and one analytical organisation beyond government has told us this has deterred them from applying for data from that environment in the future. The value of government data to provide evidence to help tackle the issues of the day cannot be realised if data sharing processes fail to enable timely access and publication of analyses.

The Integrated Data Service (IDS) continues to experience challenges. The independent review of UKSA recommends that government departments support the development of programmes like the Integrated Data Service (IDS). The review recognises the IDS as critical to enabling greater use of administrative data and improving the efficiency of statistical analysis across government and academia. It recommends that "the ONS [take] action to ensure that the purpose, scope and requirements of the IDS are clearly communicated and that the needs and concerns of departmental data owners are sufficiently understood." In recognition of the pivotal role the IDS could take in facilitating the sharing of data across government, OSR is currently considering what role it can play in supporting the development and success of the IDS. Many of the recommendations in this report are relevant to the work of the IDS, and we refer to the IDS specifically in discussion around recommendations 4, 8, 10, 11, 14 and 15.

Information Box 1: The Integrated Data Service (IDS)

The IDS is a cross-government project led by the Office for National Statistics (ONS). It builds on the ONS Secure Research Service (SRS), which has been providing secure access to de-identified, unpublished data to accredited researchers for over 15 years. The IDS is a central platform that provides access to data, analytical and visual tools in a secure multi-cloud infrastructure. It aims to be the single data analysis and dissemination platform within government by providing secure and co-ordinated access to a range of high-quality data for government analysts, devolved administrations and external accredited researchers.

In September 2023, the Integrated Data Service (IDS) achieved accreditation for data provision under the Digital Economy Act (2017), an important step forward in achieving its vision of bringing together ready-to-use data for the public good, by showing the confidence that the UK Statistics Authority has in the security and the data on the IDS. It is expected that from September 2024 the SRS will be closed to new project applications and all new projects will then begin in IDS. Data assets available through the platform are listed on the IDS website.

Office for Statistics Regulation

Summary of progress against recommendations

The following table gives a list of the recommendations from last year's report and OSR's overall assessment of whether there has been progress towards each recommendation being met. The table also shows where we have made minor changes to the original recommendations or altered the recommendations more significantly to change their focus (revised), and whether we have retained them. While most recommendations have been retained, some have been removed.

Recommendation	Assessment	Comment	Outcome
1: Social Licence	Some progress	The importance of tracking public attitudes to data sharing remains, and there has been some progress towards our recommendation. The importance of amplifying positive data sharing case studies has been highlighted. The recommendation has been amended to reflect these developments.	Retained with minor amendments
2: Guidelines and Support	Good progress	Our recommendation has been revised to "Public Engagement" and amended to reflect progress in the production of public engagement good-practice guidelines by both PEDRI and the ADR UK Learning Hub.	Revised
3: The Five Safes Framework	Some progress	We consider UKSA's commitment to a discussion on the current research accreditation framework as progress towards a review of the Five Safes. Our recommendation has been amended.	Retained with minor amendments

Office for Statistics Regulation

4: Privacy-Enhancing Technologies (PETs)	Good progress	Technical and collaborative developments by ONS and DSIT represent progress towards our recommendation, the text of which has been amended.	Retained with minor amendments
5: Data Literacy in Government	Some progress	We have identified some developments in support of our recommendation, such as last year's One Big Thing. Our recommendation has been amended to reflect the importance of data literacy at all levels of government.	Retained with minor amendments
6: Data Masterclass Content	Not met	Development of the content of the Data Masterclass is on hold, with no new cohorts being accepted for enrolment. Our recommendation has been amended and made stronger.	Retained with minor amendments
7: Arbitration Process	Not met	We recognised that our recommendation would be improved by focussing on collaboration and resolution, rather than arbitration. The recommendation has been revised to reflect this finding.	Revised
8: Career Frameworks	Not met	We have not identified any significant progress in the development of career frameworks to better support data sharing and linkage, so our recommendation has been retained.	Retained

Office for Statistics Regulation

9: Overview of Legislation	Closed	Our recommendation has been removed to reflect limited support for this recommendation among our stakeholders and the useful resources already published by third parties.	Removed
10: Broader use cases for data	Not met	Our recommendation has been amended to reflect stakeholder feedback and its link to our recommendation on a review of the Five Safes Framework.	Retained with minor amendments
11: Communication	Not met	We have not identified significant progress towards our recommendation, which has been revised to "Clarity and Communication".	Revised
12: Checklists	Closed	Our recommendation advising the creation of checklists to aid data sharing and linkage has been removed to reflect the findings of this report.	Removed
13: Transparency	Some progress	We have identified resources and initiatives that demonstrate progress in the transparency of recording and accessing government datasets. Our recommendation has been retained.	Retained
14: Funding Structure	Not met	Sufficient resourcing and funding remain a significant barrier to effective and timely	Retained with minor amendments

Office for Statistics Regulation

		data sharing and linkage, and our recommendation has been amended to reflect this.	
15: Sufficient Resources	Not met	We have not identified any specific examples where extra resource has been allocated in support of our recommendation. Our recommendation has been revised to focus on metadata and documentation.	Revised
16: Standardisation	Some progress	The importance of standardisation to support data sharing and linkage is being increasingly recognised across government, and we have retained our recommendation.	Retained

Next steps

OSR has an ongoing role in driving forward data sharing and linkage across government, beyond this follow-up report. As such, we intend to report again on the state of progress in coming years, using our refined recommendations from this review as a guide. In the meantime, OSR will convene, support, advocate and warn on specific areas where we think we can have a useful impact. In doing so, our goal is to help government, researchers and the private sector move towards a future where data sharing and linkage is widely and safely employed for the public good.

Update on recommendations

This section summarises progress against the recommendations in OSR's previous report on data sharing and linkage across government. For each theme and recommendation, we have also considered whether recommendations remain relevant and fit for purpose, and whether we can identify any gaps that we should address going forward. In compiling this information, we have relied not only on our own research, but also the extensive and valuable feedback provided by our stakeholders across government and the research community.

Public engagement and social licence

Overall, there has been good progress made against our recommendations in the areas of public engagement and social licence. Ensuring public confidence in data sharing and linkage remains crucial, and developments over the last year indicate that increasing attention is being paid to maintaining social licence to use these data as demands for data use grow. However, meaningfully engaging with the public on decisions made about using data for research and people's awareness of data as a public asset could still be improved, and public attitudes towards data sharing could be better monitored.

The Public Engagement in Data Research Initiative (PEDRI) has undertaken positive work to develop guidelines to help researchers conduct public engagement which, when published, will strengthen this important aspect of research activity. The Five Safes Framework remains a widely recognised and employed tool, helping organisations and individuals approach data sharing safely and legally. UKSA is taking forward initiatives to determine if the Five Safes can be usefully updated.

The public's key concern regarding data use remains data security. It is therefore likely that many people's attitudes towards data sharing and linkage continue to be influenced by concerns around data security, particularly when data are identifiable. A UK-wide public dialogue report produced by OSR and ADR UK showed that the public support the use of government data in research and statistics, as long as they're informed of their use and what safeguarding is in place. As such, the significant advances made in the creation and trialling of privacy-enhancing technologies (PETs) over the last year represent promising and important developments in increasing public and institutional confidence in data sharing for the benefit of all. The establishment of the Department for Science, Innovation and Technology (DSIT) in 2023 has resulted in positive initiatives to overcome the technical and policy-based barriers to data sharing, which should be maintained and enhanced.

To reflect the progress made, as well as stakeholder feedback, our recommendations on public engagement and social licence have been updated to continue to guide system-wide approaches to these areas of data sharing and linkage.

Office for Statistics Regulation

Recommendation 1: Social Licence

Original recommendation: The government needs to be aware of the public's views on data sharing and linkage and to understand existing or emerging concerns. Public surveys such as the 'Public attitudes to data and AI: Tracker survey' by the Centre for Data, Ethics and Innovation (CDEI) provide valuable insight. They should be maintained and enhanced, for example to include data linking.

Key findings

- It remains vital to engage the public to track attitudes towards data sharing so that the social licence for greater data sharing and linkage is understood and maintained.
- Wave 4 of the government's public attitudes to data and AI tracker survey (2024) will provide valuable and novel insight into public opinion surrounding data sharing.
- Clear and consistent communication supported by meaningful public engagement is key. Government and public bodies should publish and amplify positive stories of data sharing and linkage in action to inform the public of what their data has helped improve and to increase understanding of the value of data as a public asset.
- Public bodies should continue to look for opportunities to seek public feedback on specific data sharing proposals, as well as attitudes towards data use in general.

Summary of findings

The importance of having and maintaining social licence through government understanding of public views on data sharing and linkage was highlighted in our first recommendation in our last report. Through our research for this follow-up report, our findings on the role of social licence have been reinforced. While Wave 3 of the Public attitudes to data and AI tracker survey (2023) published by the Responsible Technology Adoption Unit (RTA) did not seek to measure public attitudes towards data linking specifically, it continues to provide valuable insight into how society views data use. Notably, the 2023 survey found that significant sections of the population continue to feel a lack of control over how their data are used and believe that not all groups in society benefit equally from data use. These findings align with consistent feedback from our partners around the need to involve the public in decisions made about data and the ongoing challenge of raising awareness of the public benefit of increased data sharing. But the findings also show that linking and sharing initiatives must be done transparently and with public engagement whenever possible. We also know that transparency with the public as to the use of their data is important. Moreover, we heard that there continues to be some nervousness among senior leaders regarding the potential for increased data sharing to lead to a negative public response. In addition, more can be done to promote the benefits of 'opting in' to data sharing initiatives among the public, especially in sensitive areas

such as health research. As such, we consider that our recommendation remains relevant, and have updated it to reflect stakeholder feedback.

During interviews it was suggested that government could have a role in improving public understanding of data sharing and its positive impacts, with much of the current public discourse focussed on legitimate data security concerns. Open, consistent and clear communication with the public using understandable terminology will be key to informing them of the benefits of wider data use. To aid public understanding and awareness of the value of data as a public asset, government and public bodies should publish and amplify positive stories of data sharing and linkage in action. The current strategic role of the Department for Science, Innovation and Technology (DSIT) in developing and cohering policy towards data access for the benefit of all of society and across the economy will help bring cross-government attention to the importance of social licence. There is value in ongoing conversations in government around the concept of the 'public good' in data use, and we encourage data producers and users to ensure that the public good guides their approach to all sharing and linkage initiatives.

The Office for National Statistics (ONS) has undertaken several initiatives related to social licence and public engagement over the last year. The June 2023 outreach research by ONS, What we know from engaging with the public on data, highlighted important opportunities and issues in enhancing public trust in data use and sharing. The research found that people are more likely to be at ease and consent to data linkage and sharing if it is being carried out for the benefit of society. In August 2023, the ONS also publicised the findings of the ONS Data Debate, which asked people aged 18 to 24 years old for their views on data sharing. ONS's blog post on its refreshed Data Strategy, Future-proofing our valuable commodity, reflects a transparent approach to data sharing which will help encourage public confidence. Initiatives such as these both increase awareness of data sharing in society and build public trust in the organisations which hold, analyse and share data for the public good. Public bodies must continue to look for opportunities to seek public feedback on specific data sharing proposals, as well as attitudes towards data use in general, such as through the ONS stakeholder forums (case study below) and the planned Public attitudes to data in the NHS and social care programme of public engagement. OSR also notes the value of greater research into specific cases of social licence and data linkage, such as the Parental social licence for data linkage for service intervention project funded by the Economic and Social Research Council.

Case study: ONS stakeholder forums. As the UK's largest producer of official statistics, the Office for National Statistics (ONS) is one of the most significant collectors and holders of data. Its datasets are often linked with those held by other government departments and shared with external bodies to aid research.

ONS's stakeholder forums allow these partner organisations to share insight and constructive feedback on behalf of the sectors they represent. Among the central themes of this work are public understanding and trust in data, and improving inclusivity in the collection, analysis and reporting of data. In addition, quarterly ONS Assemblies bring together strategic representatives and umbrella bodies from the charity, civil society sectors in a discussion forum to provide insight, support and critical feedback on ONS plans and approaches.

The forums represent a useful tool to help organisations maintain and enhance social licence towards data sharing and linkage, and the model could be adopted more widely across government.

OSR has an ongoing role to play in championing data sharing and linkage in the public domain beyond this report. We welcome RTA's consideration of our recommendation that future tracker surveys should include specific questions on public attitudes to sharing and linkage, and at the time of writing, a question on this topic is being developed for Wave 4 of the survey. The results of the Wave 4 tracker survey are due to be published in November 2024.

Revised Recommendation 1: Social Licence. Government needs to be aware of the public's views on data sharing and linkage, particularly for groups affected by specific projects, and to understand existing or emerging concerns. Public surveys and research such as the 'Public attitudes to data and AI: Tracker survey' by the Responsible Technology Adoption Unit (RTA) and the ONS's 'What we know from engaging with the public on data' provide valuable insight. They should be maintained and enhanced to include specific questions on attitudes towards data linkage. To improve public awareness of the benefits of data sharing, organisations should publish and amplify the positive outcomes of sharing initiatives to reinforce the concept of data as a public asset. As part of this, transparency is key to gaining public trust and so organisations should be open about potential uses for the public's data.

Recommendation 2: Guidelines and Support

Original recommendation: When teams or organisations are undertaking data sharing and linkage projects, there is a growing practice of engaging with members of the public to help identify concerns, risks and benefits. To help teams or organisations who are undertaking public engagement work, best-practice guidelines should be produced, and support made available to help plan and coordinate work. This should be produced collaboratively by organisations with experience of this work for different types of data and use cases and brought together under one partnership for ease of use. We consider that, given its current aims, the Public Engagement in Data Research Initiative (PEDRI) could be well placed to play this role.

Key findings

- The work of PEDRI in developing guidelines for public engagement activity is welcome progress towards our recommendation.
- More can be done to foster a culture in which the value and importance of public engagement in research projects is recognised.

Summary of findings

Our previous report highlighted the importance of public engagement to improve the transparency of research work and public confidence in data sharing and linkage more generally. We also reported, however, that there can be a lack of understanding about how to carry out public engagement in a meaningful way, and that this could be helped by publishing best-practice guidelines.

There has been significant and welcome progress towards our recommendation to produce guidelines for those in research and data services who undertake public engagement work. The [Public Engagement in Data Research Initiative](#) (PEDRI) has led sector-wide collaboration on developing principles for public involvement and engagement over the last year, and our recommendation has influenced the decision of some stakeholders to remain involved with the initiative. In summer 2023, PEDRI held a public consultation on the draft Best Practice Standards for Public Involvement and Engagement (PIE) in data research and statistics. This was supplemented by workshop consultations carried out from December 2023, with the [public consultation findings published](#) in April 2024. The draft standards will be refined before the publication of a final report. PEDRI also intends to create a resources hub on its website dedicated to signposting and creating new resources where gaps are identified. PEDRI aims to avoid duplicating existing efforts by concentrating specifically on data and statistics, leveraging the platform to enhance the use of current resources. These developments are extremely positive, and OSR hopes that the momentum can be maintained by future phases of PEDRI. In addition to guidelines, PEDRI should showcase exemplars of public engagement to provide role models for its engagement policies.

In addition, the Administrative Data Research UK (ADR UK) [Learning Hub](#), which contains useful resources on [Public engagement in practice](#) to guide users on PIE principles to follow, was launched in September 2023. Beyond PIE, the Learning

Hub brings together information on skills and resources for those using administrative data and data linkage and has already received positive feedback from users.

While the outlook is positive, there is still work to be done to foster a culture in which those working in data recognise the implicit value of public engagement. The publication of more supporting materials around PIE would make it easier for researchers and organisations to consistently consider public engagement requirements when planning their work. However, we also see a role for project funders in stipulating the need for PIE activity when developing projects alongside researchers.

When being transparent with the public, consideration should be given to how aspects of data sharing and linkage are described to ensure they are understandable and consistent. Following feedback, we came to recognise that guidelines describing 'good' practice are more appropriate than attempting to define 'best' practice. Standards and guidelines should therefore always be evolving, and we will make our terminology consistent with that being adopted by PEDRI. PEDRI also suggested that our recommendation could refer to 'affected groups' in addition to the public in general, to demonstrate that data access and research often impact specific demographic groups which may require tailored and sensitive approaches to engagement. This complements PEDRI's important findings on prioritising equity, diversity and inclusion in effective PIE. We have updated our recommendation accordingly.

Revised Recommendation 2: Public engagement. When teams or organisations are undertaking data sharing and linkage projects, engagement with the public and affected groups should be prioritised to help identify concerns, risks and benefits. To help teams or organisations who are undertaking public engagement work, good-practice guidelines should be produced and support made available to help plan and coordinate work. Guidance should be produced collaboratively by organisations with experience of this work for different types of data and use cases and brought together under one partnership for ease of use. We consider that, given its current aims and activity, the Public Engagement in Data Research Initiative (PEDRI) is well placed to lead this work.

Office for Statistics Regulation

Recommendation 3: The Five Safes Framework

Original recommendation: Since the Five Safes Framework was developed twenty years ago, new technologies to share and link data have been introduced and data linkage of increased complexity is occurring. As the Five Safes Framework is so widely used across data access platforms, we recommend that UK Statistics Authority review the framework to consider whether there are any elements or supporting material that could be usefully updated.

Key findings

- The Five Safes Framework remains positively recognised and widely employed by the UK research community.
- There remains an appetite for a review of the framework to determine its effectiveness in enabling safe access to data for research.
- UKSA have led a strategic workshop to discuss the current trusted research environment accreditation framework, to include the Five Safes.
- More can be done to promote the Five Safes in the UK research community.

Summary of findings

For over 20 years, the Five Safes Framework has been voluntarily adopted by data services to help them provide safe research access to data. Last year, we recommended that, owing to advancements in technology and ever-increasing data sharing requirements, the Five Safes Framework should be reviewed to consider whether there are any elements or supporting material that could be usefully updated. We considered that the UK Statistics Authority (UKSA) was best placed to lead any such review.

To date, no formal review of the Five Safes has taken place, in part owing to resourcing at UKSA. In November 2023, researchers at the University of West England published a paper describing the Five Safes and looking to the future of the framework. The framework's principles of safe data, safe projects, safe people, safe settings and safe outputs remain highly valued by users and an effective tool for ensuring the security of a data service. The model continues to be widely implemented across the UK research community. Nonetheless, there is still an appetite for a conversation about how the Five Safes could be amended or updated to reflect the needs of today's users. In February 2024, OSR published a blog post titled The success and potential evolution of the 5 Safes model of data access. Some of the views we heard from stakeholders, particularly around the concept of 'safe programmes', reflected suggestions made in the blog post. Some stakeholders, however, raised valid concerns about the risk of changing recognised definitions.

Stakeholders also told us that more could be done to operationalise the framework, in the form of improved guidance, and that the Five Safes could be better promoted among the research and analyst communities. The flexibility of the Five Safes was both lauded and raised as a concern, with an acknowledgement that organisations self-regulate their use of the framework. Good-quality and accessible guidance on the framework, such as that published by Research Data Scotland, is important in

Office for Statistics Regulation

helping data providers implement it. The [Standard Architecture for Trusted Research Environments](#) (SATRE) provides knowledge and best practices for trusted research environments (TREs), covering information governance, computing technology and data management. SATRE aims to bring together and enable the implementation of standards and frameworks that apply to TREs, including the Five Safes.

The range of views reflects the continued value in having an open discussion on the framework. As such, the UKSA's Research Accreditation Panel (RAP) arranged a strategic workshop in June 2024 to consider the effect of new technologies and concepts in the data access space. The workshop aimed to understand how these could impact the existing trusted research environment accreditation framework under the Digital Economy Act 2017. It is hoped that this initiative will prompt a community-wide conversation around the Five Safes, even if the conclusion is that the framework remains fit for purpose. OSR also recommends that, following this discussion, UKSA lead promotional work to ensure continued adoption of the Five Safes across the research community. Our report recommendation has been updated to reflect these developments.

Revised Recommendation 3: The Five Safes Framework. Since the Five Safes Framework was developed twenty years ago, new technologies to share and link data have been introduced, and data linkage of increased complexity is occurring. As the Five Safes Framework is so widely used across data access platforms, we recommend that UK Statistics Authority lead a community-wide conversation to consider whether there are any elements or supporting material that could be usefully updated.

Recommendation 4: Privacy-Enhancing Technologies (PETs)

Original recommendation: To enable wider sharing of data in a secure way, government should continue to explore the potential for Privacy-Enhancing Technologies (PETs) to be used to enhance security and protect privacy where data are personally identifiable. The Office for National Statistics (ONS) Data Science Campus is well placed to lead and coordinate this work.

Key findings

- There has been significant progress in the development and trial application of PETs over the last year.
- PETs remain developmental and expensive to design, which are barriers to their wider adoption.
- Government should continue to invest in PET research and explore their potential, coordinated by the Department for Science, Innovation and Technology (DSIT).

Summary of findings

Our previous report highlighted the benefits of Privacy-Enhancing Technologies (PETs) and recommended that government continue to explore their potential for enabling organisations to share and use people's data responsibly, lawfully and securely. These new technologies aim to mitigate the risks associated with sharing and linking datasets, particularly the risk of data breaches. We have seen significant progress in technological and collaborative work in this area over the last year. There is growing interest in PETs, with private sector investment in PET development complementing government-led initiatives.

The ONS Data Science Campus (DSC) has conducted important exploratory work on PETs in the UK. In April 2024, DSC released an experimental [Privacy Preserving Record Linkage toolkit](#), which combines secure cloud computing with innovative data linkage methods to achieve accurate linkage capabilities without sharing personal information. In the last year, DSC has also been working with the Integrated Data Service (IDS) to investigate how [synthetic data](#) can be used more effectively. In November 2023, it published a [technical report](#) setting out how it synthesised the linked 2011 Census and deaths dataset while preserving its confidentiality. These innovations demonstrate significant progress in what remains a challenging and expensive, but important, area of technological development. Although resourcing is likely to constrain further PET development by DSC in the short term, OSR recommends that ONS continue to invest in PETs and would welcome the production use of the Privacy Preserving Record Linkage toolkit in the future.

There is widespread enthusiasm for PETs and acknowledgment of the role they could play in increasing the feasibility and security of data linkage. This interest has been reflected in the Responsible Technology Adoption Unit (RTA)'s and ONS DSC's jointly run cross-government PET community of practice, which was organised to discuss case studies and risks and share ideas on potential developments. However, PETs remain developmental, with significant costs associated with research in this area and the underlying risk of data leakage

Office for Statistics Regulation

remaining, albeit reduced. These difficulties can make it difficult for organisations both within and outside government to develop business cases for the adoption of PETs. Furthermore, it was emphasised that users should not regard PETs as a silver bullet to meet all data protection requirements. Organisational controls remain as important as technical controls and data processing must be lawful, fair and transparent as well as secure.

Building on the Information Commissioner's Office (ICO) publication of guidance on PETs in June 2023, the RTA is working with ICO to develop a tool to support organisations in assessing the costs and benefits of adopting PETs. This tool will encourage potential users to consider how PETs can help them meet their data sharing requirements under a privacy-preserving federated learning method, while also highlighting the strengths and limitations of these new technologies. The coordinated approach being taken towards PETs across government is welcome, and OSR looks forward to further guidance and resources being published on technical approaches to securing data. Innovation and collaboration are also taking place around PETs in wider society, as demonstrated by the Privacy Enhancing Technologies Symposium organised by the University of Bristol and REPHRAIN and the Royal Society's Privacy Enhancing Technologies programme. Furthermore, there have been positive developments in exploring the implications and potential benefits of synthetic data, such as the funding for two projects announced by ADR UK in April 2024.

During interviews for this report, our stakeholders told us that our recommendation had directly influenced the strategic direction adopted by DSC and others in relation to PETs over the last year. Since PETs sit at the intersection of data privacy and data innovation, the DSIT RTA can play a vital role in fostering a culture of innovation, alongside partners such as the ONS. Our recommendation has been updated to reflect this. To ensure developments are acceptable to users, DSIT must involve data owners and trusted research environments in this work.

Case study: Privacy Preserving Record Linkage toolkit. The ONS DSC's Privacy Preserving Record Linkage toolkit (PPRL) is an experimental new technology which aims to help analysts link data across organisational boundaries, without creating data privacy risks. The PPRL currently consists of a Python package that implements an experimental private data linkage algorithm, a demo which shows how two organisations could use the package to perform eyes-off data linkage, and accessible tutorials to guide users in the use of secure cloud technologies. While it is only an exploratory proof of concept at this stage, the toolkit's initial results show strong performance, with an ability to match records to a high degree of accuracy. Although the toolkit is currently tested on small-to-medium sized datasets, there is the possibility to test scalability to larger datasets. The DSC hopes that the PPRL, which was developed in collaboration with NHS England, will stimulate the wider community to collaborate in the development of new privacy-preserving linkage methods. The toolkit is currently available on GitHub for anyone to test and provide feedback.

Revised Recommendation 4: Privacy-Enhancing Technologies. To enable wider sharing of data in a secure way, government should continue to explore the potential for Privacy-Enhancing Technologies (PETs) to be used to enhance security and protect privacy where data are personally identifiable. The Department for Science, Innovation and Technology is well placed to lead and coordinate this work.

Office for Statistics Regulation

People

OSR has previously highlighted that ensuring effective data sharing and linkage for statistics and research in government goes beyond technological advancements. At every step of the pathway to share and link data, the people involved, and their skills and expertise, are instrumental to projects' success or failure. A lack of awareness, priorities and capability of people involved in decision making and development, including senior leaders, analysts and those in data governance and guardianship roles, can impose barriers to successful data sharing and linkage. In our previous report, we highlighted barriers caused by lack of leadership, low risk appetites, lack of data literacy, low awareness of the benefits of data sharing and linkage, and staff retention and skills. Based on these observations, we made recommendations that the continued attendance and development of the ONS Data Science Campus's Data Masterclass for senior leaders in government be supported; that an arbitration process be used to help resolve disputes over data sharing; and that relevant analytical career frameworks be updated to ensure skills that relate to data and data linkage are consistently reflected.

During our research for this report, we found that progress relating to data literacy and skills has been slow. The ONS Data Science Campus has informed us that the development of new Data Masterclass content has been paused due to uncertainty around resource allocation. Furthermore, it is restricting enrolment in the current Masterclass to existing stakeholders. The Analysis Function (AF) and Government Digital and Data (GDAD) career frameworks have not been updated to include any data sharing and linkage-specific content, despite the AF career framework being refreshed in late 2023.

There has been growing recognition within government of the importance of data literacy across the Civil Service, not just among senior leaders. Last year, this recognition led to the One Big Thing initiative, which promoted analytical training for all civil servants. It will take sustained investment in such initiatives to ensure lasting change.

There has been progress, however, in work related to arbitration. Since our previous report, the CDDO has established the Data Sharing Network of Experts (DSNE), which helps organisations identify legal gateways to data sharing and linkage whilst also helping departments resolve questions or any differences of opinion around the appropriateness of data sharing. As such, we have updated our recommendation to draw attention to this good work and to encourage more development.

Office for Statistics Regulation

Recommendation 5: Data Literacy in Government

Original recommendation: To gain the skills to create and support a data-aware culture, it is important for senior leaders to have awareness of and exposure to data issues. One way to raise awareness and exposure would be for senior leaders to ensure that they participate in the Data Masterclass delivered by the ONS Data Science Campus in partnership with the 10 Downing Street (No10) Data Science Team.

Key findings

- A lack of data literacy and awareness of the benefits of data sharing and linkage among senior leaders continues to lead to overcaution and to be a barrier in some cases.
- The Data Masterclass, an initiative delivered by ONS Data Science Campus to strengthen data awareness and literacy among senior leaders across government, is currently restricted in terms of enrolment and content. The development of similar programs could help fill this gap in some parts of government.
- It is not just senior leaders who need awareness and understanding of the benefits of data sharing and linkage. All civil servants need some level of data literacy.
- One Big Thing is a positive example of government trying to raise data literacy across the civil service. Sustained investment in such initiatives will be needed to ensure lasting change.

Summary of findings

In our previous report, we spoke about the roles Accounting Officers, Chief Data Officers (CDOs) and Data Protection Officers (DPOs) have in progressing or hindering the adoption of data sharing and linkage for statistics and research. We found that low risk appetites and varying levels of understanding of the challenges and opportunities associated with data sharing and linkage are major barriers to adoption. Our recommendation aimed to encourage senior leaders and decision makers in data to upskill through resources like the ONS Data Science Campus's Data Masterclass.

Through conversations with stakeholders, we learnt that data literacy among senior leaders in government is still perceived as an important factor in enabling data sharing and linkage. Despite a growing focus on data literacy and [commitments from the UK Government](#) to put data more fully at the heart of government decision making, there is more to be done.

In particular, attitudes towards data sharing and perceptions of the potential risk–reward trade-off can still be a barrier to collaboration. The benefits of sharing and linking data may be widely spread, but if something goes wrong, the effects can be close to hand and potentially difficult for individual organisations. Senior leaders can come from backgrounds where little knowledge of data is required, which can cause them to be overly cautious. Greater understanding of the potential benefits and

Office for Statistics Regulation

perceived public perceptions of data sharing and linkage would enable organisational change and remove barriers. There is a leadership role for the analytical services in championing the benefits of data reuse and linkage. Developing a shared narrative on areas where linking data could really benefit decision making would be useful to help departmental leadership to see the benefits of linking data around a person, place or business.

Our previous report highlighted the Data Masterclass as an effective way of introducing senior leaders across government to the power and potential of data, and to different ways of working with data. Between July 2023 and March 2024, the Data Masterclass for Senior Leaders was started by approximately 5,000 public sector staff, bringing total enrolments since its launch in 2020 to 15,000 individuals from more than 50 public sector organisations. These data are impressive and demonstrate the reach and potential of the Masterclass. Furthermore, we heard that similar programs are being developed to support this agenda. These include the Digital Excellence Programme, which is part of the Digital and Data Civil Service Reform Missions and aimed at training SCS in data literacy and AI, and the Welsh Government internal policy essentials course for senior leaders to help improve their skills in relation to data.

Of course, it is not just senior leaders who are important in enabling data sharing and linkage. Individuals from many professions and at different levels of seniority within the Civil Service all have a role to play in understanding and promoting opportunities for sharing and linking data. This is not least because those operating at the working level are often those who brief and advise senior decision makers. Since 2020, the Data Linkage Champion Network has provided a forum for civil servants of all grades and from across government to discuss and promote data linkage. Looking more widely at data literacy across the Civil Service, in 2023 the Civil Service launched One Big Thing, a new annual initiative led by the Data Science Campus, to help civil servants upskill on topic areas relevant to a modern Civil Service. The first topic of this initiative was data skills: 212,000 people across the Civil Service took part over four months, undertaking at least one day's worth of analytical training. This course is now available for all civil servants to access on Civil Service Learning.

Initiatives like the data-focussed One Big Thing event help to promote a data-aware culture within government. This is a positive step forward, although it will take sustained investment in such initiatives to ensure lasting change. Data awareness will be essential for all civil servants given the push towards a modern Civil Service; our recommendation has been updated to reflect this.

Revised Recommendation 5: Data Literacy in Government. To gain the skills to create and support a data-aware culture, it is important for all civil servants to have awareness of and exposure to data issues. To support this, more opportunities for raising Civil Service-wide data awareness should be created. Senior leaders should continue to participate in courses like the Data Masterclass delivered by the ONS Data Science Campus in partnership with the 10 Downing Street (No10) Data Science Team.

Recommendation 6: Data Masterclass Content

Original recommendation: The Data Masterclass could expand its topics to include sections specifically on awareness of data linkage methodologies, the benefits of data sharing and linkage and awareness of different forms of data. This would fit well under the Masterclass topics of 'Communicating compelling narratives through data' or 'Data-driven decision-making and policy-making'.

Key findings

- The Data Masterclass continues to be delivered, but only to existing stakeholders.
- Development of the Data Masterclass has been paused.

Summary of findings

In our last report, we recommended that the content of the Masterclass be expanded to include more sections on awareness of data linkage methodologies. This recommendation aimed to reinforce the effectiveness of recommendation 5 through ensuring that the Data Masterclass content is fit for purpose.

From our engagement with the Data Science Campus (DCS), we found that development of the Data Masterclass has been paused while decisions on resource allocation are made within ONS. This means that, despite agreeing in principle, DSC has been unable to commit to developing the Masterclass in line with our recommendation. Furthermore, we heard that the Data Masterclass will only be delivered to existing stakeholders, reducing the number of senior leaders who would benefit.

Strengthening data literacy among senior leaders is vital to furthering data sharing and linkage, and other innovations related to data or analysis, across government. Buy-in from leaders is often required to progress positive data sharing and linkage initiatives. The Data Masterclass is an excellent tool for improving data literacy among senior leaders in central government and across the wider public sector. To continue to do this job well, the Masterclass needs to evolve to reflect developments in the data space, including data sharing and linkage. We are disappointed to hear that future investment in the development of the Masterclass is not ensured and would encourage ONS to prioritise this investment.

Given the importance of increasing awareness among senior leaders of the potential benefits of data sharing and linkage for the public good, we consider that our previous recommendation remains important and should be stricter.

Revised Recommendation 6: Data Masterclass Content. The Data Masterclass should expand its topics to include sections on awareness of technical issues including data linkage methodologies and different forms of data, as well as broader issues, including the benefits of data sharing and linkage. To aid this development, ONS should ensure that the Data Masterclass is sufficiently resourced so that it remains an excellent tool for improving data literacy across senior leaders in the public sector.

Office for Statistics Regulation

Recommendation 7: Arbitration Process

Original recommendation: To facilitate greater data sharing among organisations within government, a clear arbitration process, potentially involving ministers, should be developed for situations in which organisations cannot agree on whether data shares can or should occur. Developing such an arbitration process could be taken on by the Cabinet Office, commissioned by the Cabinet Secretary and delivered working with partners such as No10 and ONS.

Key findings

- There has been progress in this space since our previous report. The Central Digital Data Office (CDDO)'s establishment of the Data Sharing Network of Experts (DSNE) to help departments deal with questions or differences of opinion around data sharing is a positive step.
- In the year since our last report, Professor Denise Lievesley's Independent Review of the UK Statistics Authority was published. Within this report, a recommendation was made to progress work on data sharing and linkage.
- Stakeholders have raised concerns with us that any arbitration process could have difficult consequences for data owners, including decisions being made on their behalf while risk remained with them.

Summary of findings

We found in our last report that hesitancy around data sharing and linkage due to potential overcaution was and remains a major barrier. Our aim with this recommendation was to encourage the development of a process that would enable departments to work through potential disagreements over data linkage that stem from security and legal concerns. The inclusion of a third party to advise on potential risks and benefits would hopefully mitigate over-worrying and better enable productive data linkage.

In line with our 2023 recommendation, the Independent Review of the UK Statistics Authority, published in March 2024, clearly reinforced the need for better processes to resolve data sharing challenges. Within this review, Professor Lievesley recommended that "the Cabinet Office, HM Treasury and No. 10 must actively work to resolve the systemic, often cultural barriers to data sharing between departments. All government departments, particularly those who own significant amounts of data, must prioritise data sharing for statistics and research purposes". The government response to this independent review welcomed the review's focus on data sharing within government, and committed to setting out a more detailed response to this recommendation later this year. We hope the government adheres to these commitments.

The issue is not always that government departments disagree over whether to share data. It can also transpire that everyone agrees to a data share in principle, but it does not advance because it becomes encumbered by the complexities of the process, which can involve multiple steps and people. In these situations,

Office for Statistics Regulation

collaborative processes, rather than arbitration, are required to enable sharing initiatives to proceed.

Since our last report, positive steps have been taken towards establishing routes for departments to resolve questions about or conflicts over whether or how they can share data.

The Central Digital Data Office (CDDO) has established a Data Sharing Network of Experts (DSNE), which includes data protection leads and legal experts. This network helps organisations that want to share data identify legal gateways within the Digital Economy Act 2017 and helps departments deal with questions or differences of opinion around data sharing. The network also helps government departments to develop potential new data sharing objectives that can be put to the Public Service Delivery Information Sharing Review Board to expand the Digital Economy Act and enable new data sharing initiatives that would previously have been blocked by the act. The CDDO is doing helpful work in this area: guiding government departments through the legal challenges of sharing data and resolving disputes, offering the opportunity to mitigate a lot of the nervousness around data sharing and linkage, and leaving more room for optimism about future efficiencies and research. Due to the promise of DSNE, resources should be allocated to enable sustained activity in this space. However, the CDDO could make the DSNE more widely known, both within government and to the public. This would raise awareness that this service is available, as well as demonstrate transparency and offer reassurance that the process of sharing and linking data in government is rigorous. We have updated our recommendation to reflect the CDDO's work on the DSNE.

The issue of arbitration and how this can most effectively be overseen and delivered is a complex one. Our stakeholder engagement revealed a variety of views about the potential benefits and risks of introducing a third-party arbitrator to the data sharing process. Data owners need significant reassurance that any data they share will be handled correctly, and that the benefits of sharing outweigh the potential risk of security breaches. Some of the stakeholders believed an arbitration process may not be able to address these issues. Concerns were also raised around whether an arbitrator could override data owners' concerns about data security and take the control out of their hands. Stakeholders expressed worry that reputational damage could result from decisions made on their behalf. Others felt that the involvement of a third party, through an arbitration process, could be positive, as the responsibility for making decisions around data security, which would usually fall entirely on data owners, could be shared. However, with this view came the appreciation that finding an arbitrator willing to take on this level of responsibility could be difficult and require legislative changes around the accountability of data controllers.

Though we understand these concerns, overcaution around data sharing is a major barrier to realising its benefits, from both a research and operational perspective. It is important for government to continue to invest in processes that could help resolve disputes around data sharing and linkage. On reflection, collaborative processes wherein data owners and researchers could discuss terms and arrangements for data sharing with a third-party providing mediation to aid resolution may be less intimidating and more productive. To reflect this, we have updated our recommendation to focus less specifically on a formal arbitration process.

Office for Statistics Regulation

Revised Recommendation 7: Collaboration and Resolution. To facilitate greater data sharing among organisations within government, organisations should consult entities like the Data Sharing Network of Experts (DSNE) when they want to share data but need help navigating how when they cannot agree on whether data shares can or should occur. The networks that offer these services should be more visible within and beyond government. The Cabinet Office should continue to develop a dispute resolution service for cases that require discussion at the most senior levels in government, working with partners such as No10 and ONS.

Recommendation 8: Career Frameworks

Original recommendation: To enable more effective and visible support for the careers of people who work on data sharing and linkage, those responsible for existing career frameworks under which these roles can sit, such as the Digital Data and Technology (DDaT) [now known as Government Digital and Data, GDAD] career framework and the Analytical Career Framework, should ensure skills that relate to data and data linkage are consistently reflected. They should also stay engaged with analysts and professionals across government to ensure the frameworks are fit for purpose. These frameworks should be used when advertising for data and analytical roles and adopted consistently so that career progression is clear.

Key findings

- Though the Analysis Function (AF) career framework has been refreshed since our last report, it does not include specific material on data sharing and linkage.
- The Government Digital and Data (GDAD) Profession Capability Framework similarly has not been updated to include data sharing and linkage content.

Summary of findings

Research for our previous report highlighted a demand for data roles within both the public sector and the private sector. Due to the competitiveness of salaries in the private sector, the public sector is struggling to attract skilled individuals. Furthermore, we found that retention of talent is also a common problem, as individuals with specialised knowledge of processes move to other government departments, chasing higher salaries at the same grade. Based on these findings, we recommended that the most prominent career frameworks in government, the GDAD and AF career frameworks, be updated to include data sharing and linkage specific content.

Since our report, there has been little progress towards realising this recommendation. The AF career framework was last refreshed in late 2023. The AF career framework outlines core skills needed to work successfully across analytical roles and links to a complementary product (the technical learning curriculum), which sets out the wider learning opportunity for analysts and includes data sharing and linkage learning. Disappointingly, due to wider streamlining across the AF, workplans have shifted to a maintenance-only position for this financial year with no further developments to career framework or other currently scheduled products. The GDAD career framework, similarly to the AF framework, contains a breadth of information on what data skills are useful for analysts but does not contain any specific detail around data sharing and linkage. In addition, there is a lack of consistency in how these frameworks are used in government, which we highlighted in our previous report.

The work of the Cabinet Office and the development of the IDS illustrate that government statistics are moving towards an environment where data sharing and

linkage are more commonplace. Therefore, more work is needed to improve these career frameworks by including this type of content explicitly, so that analysts in government develop the appropriate skills to enable this transition. It is counterproductive to have positive initiatives in the space of data sharing and linkage without equipping analysts with the skillsets needed to make use of these advancements. We therefore consider our original recommendation is still fit for purpose.

Recommendation 8: Career Frameworks. To enable more effective and visible support for the careers of people who work on data sharing and linkage, those responsible for existing career frameworks under which these roles can sit, such as the Government Digital and Data (GDAD) career framework and the Analytical Career Framework, should ensure skills that relate to data and data linkage are consistently reflected. They should also stay engaged with analysts and professionals across government to ensure the frameworks are fit for purpose. These frameworks should be used when advertising for data and analytical roles and adopted consistently so that career progression is clear.

Processes

There has been mixed progress on our recommendations related to improving processes for applying for access to government data.

We have seen positive work in some areas. For example, UKSA and ADR UK have both published information on the Digital Economy Act (DEA, 2017) to help researchers wishing to access government data understand the relevant legislation.

Through discussions with stakeholders for this report, we have been made aware of additional process challenges. These include delays to output approvals by trusted research environments (TREs), barriers created by the requirement to publish all analyses enabled by the DEA (2017) Research Power and delays resulting from legal review of complex and non-standard data sharing agreements.

Academic and external organisation stakeholders have told us of frustrations around data owners and trusted research environments requiring review of outputs before publication and impacts on the ability to produce timely analysis because of the time this takes. Some have found output checking is taking longer than previously and that they are being asked to provide full reports when previously they were only required to share relevant sections.

The requirement to publish findings from analyses using government data has also been highlighted as a potential barrier to data sharing. Departments may not only be nervous about reputational risk from negative findings but also concerned that erroneous findings will be published from inappropriate analyses by analysts without a good understanding of the ways in which the data can and cannot be used. Government departments sharing data should provide detailed metadata and guidance on safe and unsafe uses of the data, as well as provide support to researchers throughout projects using their data.

The Pan-UK Data Governance Steering Group was established by the UK Health Data Research Alliance convened by HDR UK with the aim to simplify and streamline the governance processes surrounding data access. One priority area it has identified are delays caused by legal review of data access agreements, due to variation in contracts used across departments and their complexity. To tackle this barrier, the Steering Group produced a template for a data access agreement (DAA). The template is intended for use where data are accessed in a TRE or the NHS's Secure Data Environments (SDEs) for the purposes of research and development for the public good. The principles underpinning the DAA were developed with significant consultation, and it is optimised for data science use across the UK. The aim of the DAA is to provide a familiar structure and terminology to build trust among data owners, researchers and the public. HDR UK is coordinating this work and driving forward adoption among the network of those hosting TREs and SDEs in the UK. The creation of this resource is positive, and we encourage departments to engage with HDR UK and review the template to explore whether they could adopt it. More broadly, initiatives such as the DARE UK-supported TRE Community indicate increasing collaboration among TREs, which will aid development of best practice and support projects which span multiple TREs.

Office for Statistics Regulation

To reflect the mixed progress made, stakeholder feedback and reflection from OSR on the suitability of our previous recommendations, several of the recommendations on processes have been updated.

Office for Statistics Regulation

Recommendation 9: Overview of Legislation

Original recommendation: To help researchers understand the legislation relevant to data sharing and linkage and when it is appropriate to use each one, a single organisation in each nation should produce an overview of legislation that relates to data sharing, access and linkage, which explains when different pieces of legislation are relevant and where to find more information. This organisation does not need to be expert in all legislation but to be able to point people to those that are. The Office for Statistics Regulation (OSR) will help convene those in this space to understand more about who might be best placed to take this on.

Key findings

- UKSA and ADR UK have published resources on relevant legislation for researchers wishing to access data under the Digital Economy Act (2017) and the Statistics and Registration Service Act (2007).
- Gaps remain for researchers wishing to understand legislation relevant to health data access.

Summary of findings

Since our last report, several resources have been published that aim to improve researcher understanding of legislation relevant to data sharing and linkage. We have been told about the following resources:

- The UK Statistics Authority (UKSA) has [published an online resource](#) providing answers to frequently asked questions about the Digital Economy Act (DEA, 2017) Research power and the Statistics and Registration Service Act (SRSA, 2007). This resource briefly covers what data can be accessed via the Research power and SRSA, information about accredited processing environments and who can access data and for what purposes.
- ADR UK has launched an [online Learning Hub](#) that includes information on the DEA (2017) for researchers wishing to access administrative data under this legislation. It includes a [slide deck produced by the UKSA](#) that explains what the DEA 2017 Research power allows for and contains a visual map of the data access journey.

While these resources are positive examples of progress, gaps remain in the information available, for instance for researchers wishing to access health data. While the DEA 2017 provides a legal basis for health data access in certain circumstances, access to health data is often covered by alternative legal routes.

We continue to hear contrasting views on whether legislation is a barrier to data sharing or whether misinterpretation of the legislation and how it has been operationalised in policies and procedures, and the general nervousness around data sharing, creates a barrier. The [independent review of the UKSA](#), published in March 2024, found that while "often cited as an excuse for not sharing, the legislative framework is in fact enabling." Nonetheless, the conflation of government and

academic research requirements in primary legislation continues to be cited by some in government as a limitation on interdepartmental sharing and access conditions through the risk of reidentifying bodies corporate. A comprehensive overview of legislation as described in our original recommendation is unlikely to address this conflict.

To reflect the resources that are now available, and following the recognition that an overview of legislation as described in our original recommendation may not meet researcher needs, we have closed this recommendation. Instead, a key priority is ensuring communication and clarity around the implementation of legislation and what it means for those wanting to access data. This is discussed and reflected in our Revised Recommendation 11: Communication and Clarity.

Office for Statistics Regulation

Recommendation 10: Broader use cases for data

Original recommendation: To support re-use of data where appropriate, those creating data sharing agreements should consider whether restricting data access to a specific use case is essential or whether researchers could be allowed to explore other beneficial use cases, aiming to broaden the use case where possible.

Key findings

- Requiring narrow use cases for data continues to be a barrier to effective data sharing, and there is some support for allowing broader use cases when appropriate.
- Views continue to differ on what is permissible under current legislation, and clarity is needed about when broader use cases would be justified. We have heard examples of data access being approved for broader research themes during the COVID-19 pandemic.
- Decisions about allowing broader use cases for data should consider ethical as well as legislative and practical issues.
- Updating the Five Safes framework to change 'safe projects' to 'safe programmes' may facilitate and encourage a broadening of use cases for data by those creating data sharing agreements. The Research Accreditation Panel at UKSA committed to considering this proposal at its June 2024 strategic workshop.

Summary of findings

In our engagement for this follow-up report, we heard general support from several stakeholders for broadening use cases for data. Furthermore, requiring highly specific use cases continues to be a barrier to data access, research and the ability to use data in policymaking. Concerns that data use cases are too tightly defined to enable the use of data in policy development are particularly relevant to the success of the Integrated Data Service (IDS).

In February 2024, our Director General for Regulation, Ed Humpherson, published a blog post on [The success and potential evolution of the 5 Safes model of data access](#) discussing the five safes framework of data sharing and setting out the case for a change of focus from 'safe projects' to 'safe programmes'. This proposal aims to enable those overseeing accreditation and data access to approve access for broad areas of research rather than a tightly defined research question. The blog post recognised that benefits could include a more efficient and flexible system, where researchers would not be required to define in detail the specific variables and analysis plans that they intend to use. It would also negate the need for lengthy reapplications to answer related research questions.

Conversely, there are specific concerns about the recommendation to broaden use cases for data and the proposal to change to safe programmes, including from some stakeholders who thought approving broader use would not align with current legislation. There are concerns that safe programmes would not be in the spirit of the

security focus of the Five Safes Framework, and one stakeholder raised ethical concerns about sharing data for broad justifications. Others, however, assess that legislation does allow for broader use cases and it is instead the interpretation of legislative restraints that has led to a narrowing of use cases. Different organisations have different uses of the terms project and programmes, and in some cases, such as during the pandemic, projects with broader research questions had already been approved. These varied views demonstrate a clear and continuing lack of agreement across stakeholders in terms of what is currently allowable. Clarity is needed around whether broader use cases can be considered 'safe' and when they would be justified.

The Research Accreditation Panel at UKSA held a strategic workshop in which it considered the framework for accreditation under the DEA (2017) and discussed the proposal to change to safe programmes. As part of this work, we encourage UKSA to clarify what is legal and practical under the current legislative framework.

Adopting this recommendation has the potential to make the process of accessing data more efficient and less burdensome for researchers and data owners, and to better enable the use of government data in research and policymaking. A change to safe programmes may be a route to enabling broader use cases, highlighting the relevance of a review of the Five Safes (Recommendation 3). There are outstanding questions around the legal and ethical implications of such a change, and assurance is needed around when a broader use case is acceptable. An important consideration will be the requirement for public engagement on any broader use cases for data. We have amended this recommendation to suggest the UKSA take a lead on this in line with their work on the Five Safes Framework.

Revised Recommendation 10: Broader use cases for data. To support re-use of data where appropriate, data owners, those overseeing accreditation and access to data held in secure environments, and those creating data sharing agreements should consider whether restricting data access to a specific use case is essential or whether researchers could be allowed to explore other beneficial use cases, aiming to broaden the use case wherever possible. Given the UK Statistics Authority's commitment to consider such a change and the overlap with Recommendation 3 (The Five Safes Framework), we think it is well placed to take a lead on this proposal.

Recommendation 11: Communication

Original recommendation: To ensure data application processes are fit-for purpose and well understood, those overseeing accreditation and access to data held in secure environments should prioritise ongoing communication with users, data owners and the public to explain and refine the information required. Wherever possible, they should offer face-to-face or virtual discussions with those applying to access data early in the process, to ensure clarity around both the data required and the process to access it.

Key findings

- A lack of clarity continues around the legal basis for information required from those applying for data access, and researchers have not received sufficient assurance.
- Concerns remain that data owners and those responsible for accrediting decisions are creating unnecessary barriers to data access by requiring information beyond the legislative need.
- Data owners and those responsible for accrediting decisions should assure themselves that their policies and procedures align with legislative requirements, making changes where possible to reduce the burden on applicants.

Summary of findings

For the previous report, OSR was told that communication between those overseeing accreditation and access to data and researchers is often challenging. Researchers experience long delays in receiving feedback on applications and report a lack of clarity around why some questions are asked, such as those requiring specific details of planned statistical methodology. Communication should be prioritised to refine and explain the information required.

We continue to hear about similar difficulties with communication. Researchers using one secure research environment have reported significant increases in wait times for data access and output approval without any clear explanation for the delays. We consider that TREs can do more to communicate with researchers about wait times for data access and output approval. TREs and accreditors should consider collecting and publishing metrics on the time taken for each stage of the accreditation, output checking and data access process. This would not only aid transparency, but also help researchers to better gauge and plan timelines for their projects.

Concerns also continue to be raised about the details researchers are asked to provide when applying for research project accreditation from UKSA. We have again been told that requiring researchers to provide such detailed information on analysis plans creates a barrier, since such questions can be hard to answer before researchers have explored and developed an understanding of the data.

Office for Statistics Regulation

UKSA told us that the information requested is determined by what legislation requires, ensuring projects can be assessed against the principles and conditions in the statutory [Research Code of Practice and Accreditation Criteria](). Research project [application guidance]() from UKSA advises that methodological information is required as the suitability of analysis plans is relevant to determining whether a project has potential to serve the public interest.

Data owners and those responsible for accrediting decisions should ensure they are not creating unnecessary barriers to data access. They should be transparent about the legal basis for the information they require from those applying for data access. They should assure themselves that their policies and procedures align with legislative requirements, making changes where possible to reduce the burden of requiring specific information on applicants. We hope that the Research Accreditation Panel's strategic workshop is the start of open discussion on this topic. Those processing data requests should provide greater clarity to researchers on why certain details are requested, with reference to relevant parts of the legislation and the Research Code of Practice and Accreditation Criteria.

UKSA is working with the Integrated Data Service (IDS) to provide pop-up information to researchers completing data access applications. This is a promising step forward, and we encourage UKSA and the IDS to ensure sufficient detail is provided, with links to additional resources should a researcher wish to know more. UKSA could also update its research project application guidance to provide more clarity on why it requires information on specific elements of methodology.

We continue to see an ongoing need for improved communication, with no significant progress having been made against the recommendation. Recognising the continued lack of explanation given to researchers around why certain details are required in project applications, we have amended the recommendation to take into account a need for reflection and review by those overseeing accreditation and access to data.

Revised Recommendation 11: Clarity and Communication. To ensure data application processes are fit for purpose and well understood, those overseeing accreditation and access to data held in secure environments should prioritise ongoing communication with users, data owners and the public to explain and refine the information required. This communication should include transparency as to what the data will be used for. Those overseeing accreditation and access to data, including the UKSA, should aim to reduce the administrative burden on applicants as much as possible, assuring themselves that their policies and procedures align with legislative requirements. Wherever possible, they should offer face-to-face or virtual discussions with those applying to access data early in the process to ensure clarity around both the data required and the process to access it.

Recommendation 12: Checklists

Original recommendation: To ensure all necessary teams are involved at the outset of a data sharing and linking project, organisations should consider the use of a checklist for those initiating data sharing. The checklist should contain all contacts and teams within their organisation who need to be consulted to avoid last minute delays.

Key findings

- Government departments need clarity across their organisation about which teams handle data access requests.
- Departments should consider having a single team responsible for requesting data from other departments and for coordinating requests made to their own department.

Summary of findings

Bringing together all the relevant people and teams within an organisation to facilitate the sharing of data is challenging. In addition, researchers can struggle to find the right person to speak to about a dataset or about access processes. Often, there is a lack of clarity within government departments as to who the relevant person or team is. As such, we previously recommended the use of checklists by organisations to ensure all relevant parties are involved and informed at the outset of a data sharing and linkage project.

Through engagement for this follow-up report, we found that a related challenge for some government departments is that they sometimes receive multiple requests for similar data from different teams at one department. This demonstrates a lack of communication and coordination around data needs within organisations. Departments should consider having a mechanism for coordinating data requests across all business areas to improve efficiency and reduce demands on departments to which data requests are being made.

To better understand these challenges and how they can be tackled, OSR needs to engage more directly with data owners and researchers in the future. On reflection, our specific focus on checklists may not be useful; instead, departments should take relevant action, which may include the use of a checklist, to ensure that the right people are involved and identified from the start of a data sharing and linkage project. To reflect this recognition, we have closed this recommendation.

Recommendation 13: Transparency

Original recommendation: Every organisation within government should be transparent about how the data they hold can be accessed and the process to follow. This guidance should be presented clearly and be available in the public domain with a support inbox or service for questions relating to the process.

Key findings

- There are many examples of public bodies publishing information about what data they hold, how data can be accessed and the process to follow.
- The CDDO is developing a data marketplace that should drastically improve the discoverability of data held across government, but we would encourage the CDDO to make this resource publicly available to support external researchers.

Summary of findings

Our previous report recommended that government organisations need to be transparent about how the data they hold can be accessed and the process to follow. We have seen progress against this recommendation, with several positive examples of work in this area:

- The CDDO is developing a data marketplace to improve the discoverability of data within government. The marketplace provides a central facility allowing those within government to find out what data are held and how data can be accessed. The data marketplace supports data discovery through the adoption of consistent metadata standards. It will provide a managed catalogue of sharable resources to support departments in being able to promote data that can be shared. It will also standardise the process by which data sharing can be agreed. This resource could provide significant benefits. The CDDO should prioritise making this resource publicly available to improve transparency and ensure external researchers are also able to make use of data held by government. Furthermore, we support a long-term approach to this project in which resources are allocated to enable sustained activity.
- ADR UK has created a searchable public metadata catalogue that contains a large amount of information about the datasets held across the ADR UK partnership. The catalogue includes a webpage for each dataset with contact details and links to information on how to access the data, as well as description of the dataset, information on coverage, and for some datasets, downloadable metadata.
- HDR UK has created a searchable public metadata catalogue which contains information from over 850 different health-related datasets across the UK. As with the ADR UK catalogue, the Health Data Research Gateway

includes a webpage for each dataset with contact details and links to information on how to access the data, as well as description of the dataset, information on coverage, and for some datasets, downloadable metadata documentation. The Gateway integrates with the [Researcher Registry](#), included as a data use register (implementing the Pan-UK Data Governance Steering Group standard supporting transparency discussed below), and also implements the Data Access Request Form Standard (for data custodians who have adopted the standard).

- The [Pan-UK Data Governance Steering Group](#) was established by the UK Health Data Research Alliance convened by HDR UK with the aim to simplify and streamline data access governance processes. One of its priority areas is improving the transparency of processes for accessing health and health-relevant data for research. Transparent and clear information about the safe and secure access to and use of health data enables researchers to navigate data access processes and helps build and maintain public trust. The Steering Group co-developed and published [Transparency Standards](#) with HDR UK's Public Advisory Board (PAB) to guide good practice. These standards highlight how the principles of transparency can be met by publishing open access data use registers. With support from the Medical Research Council, in 2023, [19 organisations](#) were awarded funding to adopt the Transparency Standards. The outputs from these awards are published here: [Vol 9; Conference proceedings for UK health Data Research Alliance Transparency Showcase.](#)

- The [MoJ Data First Programme](#) is an ambitious project with the aim of unlocking the insight stored within administrative datasets across the justice system. The MoJ has published clear information on what data are available, explained how researchers can apply for access and provided contact information so queries can be directed towards relevant teams.

- [Research Data Scotland](#) runs a [Researcher Access Service](#) for those wishing to access public data in Scotland. It publishes extensive information to support this service, including a [data access overview](#) describing the stages of applying for data access – from discovering what data are available to receiving access – with links to relevant resources.

These initiatives are reassuring, but transparency around data access remains variable across government departments. Stakeholders told us that transparency has recently become a lower priority for some departments, following a push to publish data catalogues as open data several years ago.

ONS's SRS metadata catalogue is publicly available, so both the public and researchers can see the data that are being made available for research use. However, ONS is planning to decommission the SRS as the IDS is planned to take over. At present, you need to be an Accredited Researcher to access the IDS metadata catalogue. If this is still the case when the IDS takes over, it could result in a reduction in public transparency about how data are being used.

Office for Statistics Regulation

Stakeholders informed us that some departments are hesitant to make information public about what data they hold, as this may have implications for resources and reputational risk. For instance, departments may be concerned about transparency leading to increased requests for data and data removal. Conversely, many departments are happy to have information on their data made public. Departments who have positive experiences of being transparent about data they hold should share their learnings across government.

We continue to encourage all departments to be open about what data they hold and how they can be accessed. Departments should explicitly state on their statistical webpages what data they hold and what process should be followed to gain access to them.

Following feedback, we recognise that some government organisations may themselves have a poor understanding of the data they hold and have updated the recommendation to reflect this.

Revised Recommendation 13: Transparency. Every organisation within government should be transparent about what data it holds, their potential uses, how these data can be accessed and the process to follow. This guidance should be presented clearly and be available in the public domain with a support inbox or service for questions relating to the process. Given its work developing a data marketplace, we consider the CDDO to be well placed to take a lead on encouraging and supporting organisations to implement these recommendations.

Recommendation 14: Funding Structure

Original recommendation: To allow every organisation a consistent funding stream for their projects, a centralised government funding structure for data collaboration projects across government, such as the Shared Outcomes Fund, should be maintained and expanded.

Key findings

- Sufficient resourcing remains a barrier to efficient, timely and cost-effective data sharing by government departments.
- A centralised government funding structure for data collaboration projects across government would benefit system-wide approaches to sharing.
- Investment focussed on access-based developments, as well as specific sharing initiatives, will aid a sustainable approach to collaboration.

Summary of findings

Our last report recommended a centralised government funding structure to enable greater opportunities for data collaboration projects across government. Specific initiatives are improving the culture towards data sharing, successful funding remains highly dependent on the priorities and vision within each department, and that resourcing remains problematic. Expensive IT systems, which often need to be developed to enable specific linkage projects, can be a barrier to cost-effective collaboration. Preparing datasets for sharing can be costly, time consuming and resource intensive. As part of the ongoing progress of data sharing and linkage in government, we expect departments holding key data assets to be sufficiently funded to provide these services, without burdening those who require access to data. Hence, we are updating our recommendation to include a call for additional resources in more areas related to data sharing and linkage. We also consider that the requirement for an enhanced centralised funding structure, likely coordinated through DSIT, should remain. Finally, we see an opportunity for the IDS, as a cross-government service, to help overcome the costly technological constraints on linkage experienced by individual departments.

Our previous report recommended maintaining the [Shared Outcomes Fund](), which had funded programmes such as the Ministry of Justice's [Better Outcomes through Linked Data (BOLD)]() initiative. We were pleased to see that funding has been allocated to a cross-government linkage project in Round 3, as announced in November 2023. The Refugee Integration Outcomes initiative will create an anonymised dataset of refugee integration outcomes based on linking Home Office refugee data to census and cross-government administrative data.

However, funding individual projects in isolation could deepen fragmentation in the system. As such, in addition to investing in collaborative projects, senior leaders across government should look for opportunities to fund strategic system-level development in their areas of responsibility. By prioritising an access-based approach that focuses investment in technical development (including IT systems), data cataloguing, upskilling and public engagement, departments would be better placed to then engage in specific data sharing initiatives in the future in a more

Office for Statistics Regulation

sustainable manner. Political leadership grounded in a comprehensive understanding of the benefits and requirements of sharing and linking data is crucial to achieving a consistent and sustained funding stream.

Case study: Better Outcomes through Linked Data (BOLD). BOLD, which is led by the Ministry of Justice (MoJ), is a data-linking programme which aims to improve the connectedness of government data in England and Wales. The programme was created to demonstrate how people with complex needs can be better supported by linking and improving the government data held on them in a safe and secure way. A cross-government system-level development, BOLD uses pseudonymised data from the Ministry of Justice, Department of Health and Social Care, the Department of Levelling Up, Housing and Communities, Public Health Wales and the Welsh Government. Pseudonymisation is a technique that replaces or removes information in a dataset that identifies an individual. Privacy, legal requirements and robust ethical standards are at the heart of BOLD's design and ethos. BOLD consists of four data and analysis pilot projects: reducing homelessness, supporting victims of crime, reducing substance misuse and reducing reoffending. Funding for the initiative was provided by the Shared Outcomes Fund.

Revised Recommendation 14: Funding Structure and Resourcing. To allow every organisation a consistent and sustainable funding stream for their projects, a centralised government funding structure for data collaboration projects across government should be established. This structure should prioritise a system-level, access-based approach to investment, as well as continue and expand initiatives such as the Shared Outcomes Fund. Senior leaders should ensure there are sufficient resources allocated to developing data sharing and linkage capabilities in their own departments.

Technical Challenges

In our previous report, we highlighted that technical challenges can pose significant barriers to effective and efficient data sharing and linkage. These include problems presented by variation in data standards and definitions, the effectiveness of data linkage methodologies and the recording of accurate metadata. As such, we made recommendations advocating the allocation of sufficient resources to allow the development of quality metadata and documentation, and in support of government doing more to standardise data and definitions. We also advised that the development and deployment of data linkage methodologies should be done in the spirit of transparency, openness and collaboration.

During engagement with stakeholders for this report, we heard that there has been welcome progress against our recommendations. Leadership by CDDO, particularly in creating the Data Marketplace and guiding departments in identifying and recording their Essential Shared Data Assets (ESDAs), is helping to create conditions for a government-wide approach to overcoming technical challenges. The Department for Science, Innovation and Technology (DSIT) has continued its work to map current data standards across government and has engaged those outside government to discover what types of data they would like to access and in what form. Nonetheless, significant technical challenges remain for data sharing, and consistent buy-in and proactivity by all government departments will be required to overcome them. This includes the allocation of sufficient resources by senior leaders, as this would ultimately result in more efficient and cost-effective data sharing and linkage across government.

Recommendation 15: Sufficient Resources

Original recommendation: To enable effective, efficient, and good quality data linking across government, senior leaders should ensure there are sufficient resources allocated to developing quality metadata and documentation for data held within their organisations.

Key findings

- The CDDO is leading efforts to help departments identify their Essential Shared Data Assets (ESDAs), which will be crucial to improving the documentation of data held by government.
- More needs to be done within departments to record accurate metadata for ESDAs to ensure efficient and effective data sharing is possible.

Summary of findings

Since our previous report, we have not identified any specific examples where extra resource has been allocated to the maintenance of data documentation and metadata. However, we would like to acknowledge the CDDO's systemic work in this space, as we believe it is one of the biggest drivers for prioritising data sharing and linkage resource in government departments. The CDDO's work helping government departments identify their Essential Shared Data Assets (ESDAs) and its development and trial of the Data Ownership Model have been successful in increasing the identification and documentation of data in government. This includes the creation and maintenance of accurate metadata. These two initiatives underpin the development of the Data Marketplace, outlined in CDDO's roadmap for 2022 to 2025. The CDDO is working with analysts contributing to the IDS to use this initiative to drive the discoverability of these ESDAs and their potential availability on the IDS platform.

The CDDO defines ESDAs as "data assets that are critical from a cross-government perspective". The CDDO aims to help government departments identify these data assets within their organisations and via the Data Marketplace improve their discoverability and further potential sharing across government to improve public service delivery, analyse the effectiveness of policies and programmes and ensure the effective use of resources. In tandem with the identification of ESDAs, the CDDO is also trialling the data ownership model that is applied to ESDAs. The Data Ownership Model defines specific roles in the management and care of ESDAs. The key objectives for the Data Ownership Model include considering "where data assets may have value to wider government, society and the economy, and the protection and exploitation approaches required to realise it" and ensuring "every Essential Shared Data Asset has accurate metadata". It is our hope that the CDDO's work will both educate senior leaders on the importance of data documentation and metadata and give them the provision to recruit more analysts to take on these responsibilities.

Office for Statistics Regulation

Revised Recommendation 15: Metadata and Documentation. To support effective, efficient and good-quality data linking across government, senior leaders should ensure that quality metadata and documentation for data held within their organisations are developed. This will improve the efficiency of data sharing and linkage to enable the swifter delivery of public services and policy decisions.

Office for Statistics Regulation

Recommendation 16: Standardisation

Original recommendation: Many departments are looking to standardise government data and definitions, but it is unclear whether or how these initiatives are working together. Those working to standardise the adoption of consistent data standards across government should come together to agree, in as much as is possible for the data in question, one approach to standardisation which is clear and transparent. Given the work done by the Data Standards Authority, led by the Central Digital and Data Office (CDDO), the CDDO may be best placed to bring this work together.

Key findings

- The importance of standardisation is being increasingly recognised across government, with recent initiatives by the CDDO, ONS and DSIT taking welcome action.
- Variation in data standards and definitions remains, and departments can do more to support government-wide standardisation work to improve the potential for high-quality linked outputs.

Summary of findings

Much progress has been made in line with this recommendation since our last report. The CDDO, through the Data Standards Authority, has begun working with ONS and other departments to define data primitives, which it is defining as common data attributes shared between different departments. This work is closely linked to ongoing work with representatives from local government around data sharing needs around the identification and provision of services and support to vulnerable people. However, there is limited resource within the CDDO to carry out this work, and getting buy-in from some departments is difficult.

The standardisation of data systems is still an issue, with multiple stakeholders saying that legacy systems and unique software are still causing issues. This was raised as a problem mainly when sharing data between academia and government. Academics can often use software that requires subscriptions, whereas government favours open-source software. This creates issues as there is little desire within government to go back to using licenced software.

DSIT been working with the CDDO on addressing coherence in data sharing and linkage. The CDDO is focused on the technical aspects, whereas DSIT focuses on policy. DSIT has engaged with businesses and researchers to find out what types of data they would like to access, and in what form. Similarly to the CDDO, DSIT also wants to help government understand what data it holds and whether they are in an accessible form. DSIT told us that coherence of policy approach is needed to enable data sharing and linkage, but that consistency is not necessary for each sharing scenario. Each one will be different, and individual considerations will be needed for different datasets and approaches. However, consistency for interoperability purposes, such as through standards, remain important. DSIT is also helping the CDDO develop the Data Marketplace.

Office for Statistics Regulation

We believe this recommendation is fit for purpose.

Recommendation 16: Standardisation. Many departments are looking to standardise government data and definitions, but it is unclear whether or how these initiatives are working together. Those working to standardise the adoption of consistent data standards across government should come together to agree, as much as possible for the data in question, one approach to standardisation which is clear and transparent. Given the work done by the Data Standards Authority led by the Central Digital and Data Office (CDDO), the CDDO is best placed to bring this work together.

Office for Statistics Regulation

Annex A

Organisation and teams that contributed to this review

We would like to acknowledge the significant input of all the individuals and organisations that contributed to this review as we gathered information and tested the ideas presented within it.

Over the course of the review, the Office for Statistics Regulation gathered information, and received feedback on our analysis and recommendations, from the organisations and teams listed below. In addition to those listed below, we also discussed the report with several academics and independent researchers, each of whom has expertise and interest in the development of data sharing and linkage, who generously gave their time.

- Administrative Data Research UK (ADR UK)
- Cabinet Office: Central Digital and Data Office (CDDO)
- Data and Analytics Research Environments UK (DARE UK)
- Health Data Research UK (HDR UK)
- Department for the Economy, Northern Ireland
- Department for Environment, Food and Rural Affairs
- Department of Health, Northern Ireland
- Department for Science, Innovation and Technology (DSIT): Responsible Technology Adoption Unit (RTA), and Strategic Data Policy
- His Majesty's Revenue and Customs (HMRC)
- Information Commissioner's Office (ICO)
- Ministry of Justice
- Northern Ireland Statistics and Research Agency (NISRA)
- Office for National Statistics (ONS): Analysis Function, Data Science Campus, and Integrated Data Service
- Public Engagement in Data Research Initiative (PEDRI)
- Research Data Scotland (RDS)
- UK Statistics Authority (UKSA)
- University of Exeter
- University of the West of England (UWE)
- Welsh Government