Public opinions on accessing public sector data for research in the public good

Findings from deliberative workshops

Ipsos Scotland, August 2025

Ipsos

Research
Data
Scotland

Contents

Executive Summary

Background, aims and objectives

Research Data Scotland (RDS) is an independent charity, established in 2021 and initially funded by the Scottish Government. RDS works with researchers, analysts and policymakers to support evidence-based policy making by making it faster and simpler for approved researchers to securely access public sector data for research in the public good. RDS launched the Researcher Access Service (RAS) in April 2024 to provide secure access to deidentified person level public sector data for approved researchers.

The RAS has initially focused on health datasets, with plans to expand into areas like education and justice. At present access through the RAS has been limited to UK-based public sector bodies, but future expansion may include the private and third sectors. RDS is also establishing a Public Impact Advisory Group (PIAG) to integrate public perspectives into the data access approval process.

On behalf of the Scottish Government, RDS is developing a common Operational Framework for simplified access to public sector data by the private sector, in line with the highest ethical and legal standards. This framework will align with Scottish Government policy in this area and support the aims of the Unlocking the Value of Data programme. The Operational Framework was known as the Private Sector Access Framework at the time of fieldwork which is reflected in the language used in the main body of the report.

To ensure the development of the RAS, PIAG, and the framework for private sector access is informed by public views, RDS commissioned Ipsos to undertake deliberative research. The specific objectives of this research were to:

- Refine the definition of public good and develop criteria for assessing when access to data is for the public good to reflect access to non-health datasets.
- Explore what public involvement via the PIAG should look like.
- Explore different considerations for how private sector organisations could access public sector data.
- Capture any learning around how best to communicate concepts relating to research data projects to the public more broadly.

Methodology

This deliberative research was undertaken with 32 members of the public from across Scotland, who were recruited to be broadly reflective of the national population. They met across three workshops between April and May 2025. The process began with an online learning session to familiarise participants with RDS's work, introduce key concepts, and present the overarching questions to be addressed. There were:

- How can we tell when the use of deidentified data about people in Scotland for research is in the 'public good', and when it is not?
- How should the public be involved in decisions about sharing deidentified data about people in Scotland?
- Which methods for private sector organisations accessing data about people in Scotland are acceptable?[1]

Participants split into in-person and online groups for a second workshop to deliberate on the key questions. They then reconvened for a final online workshop to review the emerging findings, formulate their collective conclusions and consider how organisations like RDS could communicate with the public. Discussions were guided by stimulus materials and presentations from specialist speakers, who were available to answer questions throughout the process to ensure participants could form a considered view.

Key findings

How can we tell when the use of deidentified data about people living in Scotland for research is in the 'public good', and when it is not?

There was broad public support for all organisations being required to demonstrate how their proposed research serves the public good before being granted access to public sector data. Participants acknowledged the difficulty in creating a single definition of "public good" but established overarching criteria for its assessment. These include the need for tangible benefits, a clear prioritisation of public wellbeing over any financial interests, and robust safeguards to protect vulnerable groups. Participants' full conclusions are presented below (with those they highlighted as particularly important denoted with an asterisk):

Use of data is in the public good when:

***The research can demonstrate that it will result in a better understanding of an issue by addressing a knowledge gap or providing valuable new insights or demonstrating a specific and measurable benefit for the public.** Some examples might be:

- More affordable housing in the right places

---

[1] Under this question, participants were asked to share their views on three possible models for private sector access. Participants were not asked to compare or rank the models, but rather to explore the key considerations for each.

- A fairer and more supportive education system
- Support programmes to promote positive life outcomes for young offenders and help break the cycle of reoffending
- Contributing to Scotland's effort to reach net zero
- Better understanding of the needs of vulnerable groups, like young offenders or children in kinship care

The intended benefits are realistic, actionable and implementable. Where the research is exploratory and is intended to be a stepping stone to future public good, there is a clear roadmap where longer-term potential for public good realisation and intended public benefit are defined, even if not leading directly from the research, for example:

- Identifying areas where affordable homes are needed, and building them
- Improving mental health services for young people
- Targeted interventions in schools for children and young people who need it

***The research prioritises the wellbeing of people in Scotland, and the financial interests of companies are a minor consideration.** For example through:

- Reducing energy bills for households
- Improving the condition of homes (and subsequently the health of those living in them)
- Improving support to children and young people in education
- Housing developments that boost local economies and provide community benefits

It is clear who benefits, whether it is a small group or a broad range of people, and the benefits are accessible to those groups.

***The risks of the research are proportional to and outweighed by the potential benefits, and appropriate safeguards are in place to protect vulnerable populations.**

The research maintains public trust and confidence, with data being used responsibly and ethically and original research objectives are adhered to.

Where measurable benefits are expected, timescales for realising these are defined.

Use of data is not in the public good when:

The research objectives are intentionally vague, or the intended public good is unclear and hard to measure.

***There is a risk of doing more harm than good,** for example:

- Singling out, marginalising or stigmatising certain groups, such as low-income households, vulnerable people, or young offenders

- Increasing energy bills of vulnerable households

***There is potential to target vulnerable people or "make money from suffering",** for example:

- Targeting financial products at vulnerable households
- Reducing or removing benefits that people rely on

The intended outcomes OR the roadmap to potential future benefits seem unrealistic or not likely to work in practice. For example, advertising services or products (like installing a more energy efficient heating system) that cannot be delivered due to a lack of supply chain or infrastructure

Profit seems to be the main motivation, rather than a genuine desire to benefit the public, for example:

- Housing developments that are not affordable or do not address the needs of the community
- Higher profits for energy companies with no financial benefits to customers

The research does not lead to the creation of new knowledge or a better understanding of an issue.

How should the public be involved in decisions about sharing deidentified data about people in Scotland?

There was widespread support for a Public Impact Advisory Group (PIAG) as a means of embedding public involvement in the Researcher Access Service, thereby increasing trust and transparency. For the PIAG to be effective, participants concluded it must have diverse representation from a cross-section of society, be recruited through a robust and open process, and provide members with training and financial compensation to ensure broad participation. Participants' full conclusions are presented below:

- Opportunities to join the PIAG should be widely advertised to ensure fairness and representativeness. Advertising could be done via social media, in newspapers, or in public places. RDS should clearly explain what the group is and what membership will involve.
- Diverse representation is essential. The Group must include people with diverse demographics, socioeconomic backgrounds, lived experiences, and expertise. Members could be asked some screening questions before they are invited to join the Group to ensure there is diverse representation.
- Any potential conflicts of interests should be disclosed by members. Members could be anonymous to researchers and organisations to prevent potential influence by them.

- There must be clear guidelines, training and support for members of the Group so they have the knowledge they need. The content of group discussions should be confidential as members will have access to privileged information.
- Skilled and independent facilitation of the Group will be necessary to ensure all perspectives are considered in group discussions.
- Group members should be compensated financially for their time to ensure a cross-section of society can take part.
- Ideally the Group should review all applications. If this is not possible, involvement will depend on the project type and risk. All applications from the private sector should be reviewed by the Group. Even if the Group does not carry out a full review, it must be aware of all applications that are made.
- The Group could be involved at an early stage before proposals are formally submitted as a way to shape research design, identify potential risks, and ensure public good from the outset. The public could also be involved in monitoring to make sure researchers do what they say they will.
- If the RDS Approval Panel reaches a different decision to the conclusions reached by the Group, the reasons for this should be clearly explained to the Group.
- There is a need for greater transparency and public awareness about data sharing and research more generally.

Which methods for private sector organisations accessing data about people in Scotland are acceptable?

Despite some scepticism towards commercial entities, participants accepted that private sector access to public data could be beneficial if subject to the same rigorous "public good" criteria and appropriate safeguards. Vetting any private sector organisation requesting access was deemed essential, including its structure, history, values, and commitment to public good. Participants explored three models of private sector access and identified key considerations under each. Their conclusions are presented below:

Key considerations for model one (research by a private sector organisation):

- This could be seen as the highest risk model due to concerns about the potential for exploitation of data for profit, lack of impartial oversight, potential for bias in research design and interpretation, and that some private sector organisations might not have the same level of expertise in data interpretation as the public sector.
- But with sufficient safeguards in place, including input from the PIAG, this may be the fastest route to public good realisation if the research is conducted more quickly than the other models, as the public sector might not have the resources to do the same research as efficiently.

- There should be a clear requirement for the private company to make some information about the research publicly available, to avoid the risk of them 'cherry –picking' the findings that align with their commercial motives. At a minimum, this should include information about the purpose and aims of the research, progress towards aims and some findings. RDS should be able to check back against the stated aims to assess whether the company has achieved them or not.

Key considerations for model two (research by a private sector company working with public sector and others):

- Offers a balance between accountability and efficiency.
- Offers checks and balances leading to trust in adherence to the research objectives and public good realisation.
- The collaborative approach has the potential to strengthen the research through combining resources and expertise which can lead to greater public good.
- But there is potential for bias and undue influence with private companies 'cherry-picking' partners who are likely to support their agenda.
- Where a partnership is in place, the organisation funding the research should be disclosed.
- Clear roles and responsibilities of each partner should be defined at the outset.

Key considerations for model three (research for, but not by, a private sector company):

- This may not mitigate risks around private sector access if the private company is funding the research.
- It is potentially the least efficient model resulting in slower research process and public good realisation.
- Safeguards should be put in place to ensure researcher independence and prevent against undue influence from the private sector.
- There should be a clear requirement for the private company to make some information about the research publicly available.

Communicating with the general public

The research highlighted a need for proactive public communication to build awareness and trust. Key recommendations included using clear, jargon-free language across various media channels, utilising real-world examples to explain complex concepts, and promoting data literacy to the wider public.

Introduction

Background

The use of public sector data for research presents complex ethical considerations, balancing the potential for societal benefit with the need to protect individual privacy and maintain public trust.

The Scottish Government is committed to using data in ethical, transparent and trustworthy ways to deliver better outcomes for the people of Scotland. This commitment is enshrined in the National Digital Strategy for Scotland, the Health and Social Care Data Strategy and its vision for AI in Scotlandhttps://www.gov.scot/publications/scotlands-ai-strategy-trustworthy-ethical-inclusive/. It informs the Scottish Government's Unlocking the Value of Data programme, which aims to aid ethical decision-making by public sector data controllers, regarding access to public sector personal data by the private sector. Unlocking the value of this data in secure, ethical and publicly-trusted ways, has the potential to deliver significant social and economic benefits.

Research Data Scotland (RDS) is an independent charity, established in 2021 and initially funded by the Scottish Government. RDS works with researchers, analysts and policymakers to support evidence-based policy making by making it faster and simpler to access public sector data for research. RDS is developing and expanding its Researcher Access Service (RAS), which was launched in April 2024, to provide secure access to deidentified public sector data for approved researchers held in Scotland's National Safe Haven – A Trusted Research Environment. Projects are subject to review by an Approval Panel as part of the Five Safes framework, and to ensure the proposed research is in the public good. Currently, the RAS provides access to nine health and vital events (e.g. births and deaths) datasets but may expand to include other areas like education, social work and justice. The RAS currently operates with a set of public good criteria in relation to health research projects, but the ambition is to expand these for a wider range of data sets.

Applications for access to data are currently reviewed by the RAS Approval Panel formed of experts from RDS and Public Health Scotland. RDS also jointly run Scotland Talks Data with the Scottish Centre for Administrative Data Research (SCADR). This is a public panel on the use of data for research, however there is currently no public involvement in reviewing researcher requests for data through the RAS route. RDS is therefore creating a new Public Impact Advisory Group (PIAG) as part of the RAS.

The RAS is currently limited to approved UK-based public sector organisations, namely, select universities, the NHS, Local Authorities and the Scottish Government. However, access may expand to the private and third sectors in future. Given the complexities around private sector[2] access to public sector individual data, RDS is working with Scottish Government to develop an

---

[2] For the purposes of this research, private sector was defined as companies which aim to make profit.

operational framework that ensures a consistent and trusted process for enabling private sector access.

These initiatives have been informed by existing public engagement work,[3] including ongoing feedback provided by the Scotland Talks Data public panel. Previous research to explore public opinion on access to public sector data has already demonstrated that there is broad support for the use of data for research that informs public policy. However, this support is conditional on a range of considerations. Key among them:

- **Public benefit:** the importance of any use of public sector data for research having a clear purpose and a demonstrable public benefit has been consistently highlighted. There has not been one broadly accepted definition of public benefit, but desirable outcomes have included improved health, better public services and economic and environmental benefits.
- **Data security and governance:** robust security measures are considered essential, and the existence of Trusted Research Environments (TREs) reassure members of the public that data is being accessed safely. However, concerns have also been raised about the risk of data breaches, particularly when personally identifiable information is contained in the data.
- **Clear lines of accountability, communication with the public, and access by trusted organisations** (like the NHS and universities) were also felt to be important. Private sector access to public sector data for research has therefore been met with some scepticism, based on the perception that profit is the main motivation rather than public benefit.

Aims and objectives

Building on this existing public engagement work (mentioned above), RDS commissioned Ipsos to undertake deliberative research with members of the public in Scotland. A deliberative approach was undertaken to ensure that decisions around the RAS, PIAG, and Operational Framework[4] could be informed by the public as they are being developed. Deliberative research also provides participants time to learn about and discuss the issues before coming to a considered view.

The specific objectives of the deliberative research were to:

- Refine the definition of public good and develop criteria for assessing when access to data is for the public good to reflect access to non-health datasets.

---

[3] Public dialogue on data sharing outside of the public sector in Scotland – 2024 Report for Scottish Government; Public dialogue on the use of data by the public sector in Scotland – 2024 Report for Scottish Government; Public perspectives on access to health data by non-traditional researchers: findings from deliberative workshops – 2022 Report for DataLoch; The One-Way Mirror: Public attitudes to commercial access to health data – 2016 Report for the Wellcome Trust; Public Acceptability of Data Sharing Between the Public, Private and Third Sectors for Research Purposes – 2013 Report for Scottish Government.

[4] The Operational Framework was known as the Public Sector Access Framework at the time of fieldwork which is reflected in the language used in the remainder of the report.

- Explore what public involvement via the PIAG should look like.
- Explore different considerations for how private sector organisations could access public sector data.
- Capture any learning around how best to communicate concepts relating to research data projects to the public more broadly.

Methodology overview

For this research, a group of 32 people living across Scotland met across three workshops in April and May 2025. The first workshop comprised a learning session with all participants and took place online. The learning session introduced participants to the topic of deidentified public sector data for research and related concepts. This was followed by two separate workshops; one taking place in-person in central Edinburgh, and another taking place online for those living across Scotland. In these sessions participants considered the topics in more detail before starting to form conclusions on three overarching questions:

1. How can we tell when the use of deidentified data about people in Scotland for research is in the 'public good', and when it is not?
2. How should the public be involved in decisions about sharing deidentified data about people in Scotland?
3. Which methods for private sector organisations accessing deidentified data about people in Scotland are acceptable?

Participants reconvened for a final online workshop, where draft conclusions based on emerging findings (developed by the Ipsos research team) were presented to them. They heard reflections on the emerging findings from RDS before discussing and reaching final conclusions. Participants also considered how organisations like RDS could communicate with the public (appendix A).

Overall, 36 participants were recruited to take part by telephone, using a screening questionnaire (see appendix B). The questionnaire captured demographic information about the participants, designed to help ensure the group's profile was broadly reflective of the Scottish population but also that those taking part in the in-person session could travel to Edinburgh. The aim was to achieve a sample of 36 participants, accounting for potential cancellations or drop-outs. In the end, 32 participants attended all the workshops. A more detailed overview of the recruitment can be found in appendix B.

The workshops were structured around discussion guides and stimulus materials which were developed by Ipsos and approved by RDS (appendix C). Presentations to support learning were developed by specialist speakers (with support from Ipsos and RDS) and these were delivered live during plenary sessions (appendix A). The specialists were available to answer questions from participants in both the online and in-person sessions, while questions not answered in the sessions were addressed via a Q&A document and shared with participants by email (appendix D).

Analysis and interpretation

The key output from any deliberative process is the conclusions that the participants themselves reach in response to questions posed to them at the beginning. These are presented at the end of each chapter. The deliberative nature of the project allowed for ongoing analysis throughout fieldwork, which ensured that emerging themes from the workshops could be played back as deliberations progressed and form the basis of participants' conclusions. These emerging themes, showing how participants reached their conclusions, is summarised within each chapter.

The conclusions set out and discussed in this report are intended to inform the development of the RAS, PIAG and Operational Framework. The deliberative workshops supported participants to learn about how deidentified public sector data is accessed for research and to share their views on when research is for public good, how to involve the public, and the key considerations for private sector access. This report synthesises the diverse expressions of participants to draw out key themes of discussions and highlights the ways in which participants made sense of a complex topic, describing what mattered to them and why.

The report refers to verbatim assertions by participants and their understanding of the issues. These are not intended as authoritative statements of fact, but they tell us something valuable about how key concepts, activities, or approaches have been perceived and understood by members of the public.

Further, it should be noted that whilst the method of qualitative analysis is systematic and rigorous and the conclusions robust (being based on groups that are reflective of the diversity of the wider public), the analysis does not seek to quantify findings, nor does it indicate statistical significance from a representative sample.

Participants' starting points

This chapter explores participants' early views on the key concepts and reflections on information presented to them during the first learning session.

Participants initially heard how data about individuals is generated in the day-to-day work of public sector organisations such as the NHS, government departments and local authorities. The measures in place to facilitate safe and responsible use of public sector data were described, including that:

- Public sector data is de-identified to protect individual privacy, meaning that names, addresses, and other direct identifiers are removed before researchers can access it.
- Access to deidentified data is granted within Trusted Research Environments (TREs), also known as Safe Havens. Data cannot be removed from the TRE.
- Access to data must satisfy the requirements of the Five Safes framework: safe projects, safe people, safe settings, safe data and safe outputs.
- Projects are only approved if the research would be in the public good, meaning that potential benefits to the public outweigh any risks.

Having learned about the use of public sector data for research in a general sense, participants were then introduced to RDS and its work, including:

- **The Researcher Access Service (RAS),** which aims to make it faster and simpler for researchers to access public sector data securely for research in the public good. Participants heard that RDS does not currently allow private sector organisations to access public sector data, however access is allowed by other bodies.
- **The Public Impact Advisory Group (PIAG),** being developed by RDS to ensure public views are considered within the RAS.
- **Private Sector Access Framework:** commissioned by the Scottish Government, RDS is developing a framework to help ensure there is a consistent approach for private sector access.

Familiarity with data and its use for research

At the outset, there were varying levels of familiarity with the concept of public sector data and its use for research purposes. Participants commonly associated data with terms such as “security”, “privacy”, “protection”, “misuse”, “manipulation” and “data breach”. More positive associations included “power”, “information”, “value”, “resource” and “future” and were commonly mentioned by those with more familiarity on the topic through their work. Overall, the use of public sector data for research was a relatively new topic.

Initial views were dominated by a sense of concern about personal data being used unethically and for profit, data falling into the wrong hands and breaches of privacy and confidentiality.

> "I think that data can be used and publicised in ways that can manipulate people's opinions. Can be used by companies to sort of get what they want. So I think it's something that needs to be very carefully monitored." (Participant, Workshop 1)

Potential benefits were also recognised, such as the value in using public sector data to plan for the future, improve services and understand more about public opinion on certain issues.

> "Information gives a picture. I think of things like the census, where people live, how people move, work, their health – that can shape services potentially. So, I do see that there's benefit to it by understanding, getting a quantity of information that you can analyse. But I think, manipulation came up a few times in the chat and I think that is at the front of my mind – how information can be used negatively." (Participant, Workshop 1)

Reflections on presentations and key concepts

Presentation 1: use of public sector data in research

Delivered by an expert from DataLoch, the presentation covered what is meant by public sector data, deidentified person-level data and how data are protected and used for research. It introduced the Five Safes Framework and Trusted Research Environments (TREs) or Safe Heavens and discussed the risks and benefits of using such data for research.

The presentation elicited positive reactions overall. Some found the examples of how data has been used in practice "fascinating" and helpful to understand the value of using deidentified public sector data for research.

> "It [the presentation] showed me a good way of using our data [addressing] my concerns about privacy and things. So, if it can be used for good, then, yes, I'm all for that. There were some good examples there." (Participant, Workshop 1)

Having learned more about how public sector data are protected, there was a sense of reassurance by the levels of scrutiny and security. The extent of the measures was surprising to some and de-identifying data was seen as a particularly important safeguard. Questions were raised over whether it would still be possible to trace data back to individuals once deidentified, especially if multiple datasets were linked. This concern persisted throughout the sessions.

The Five Safes Framework was viewed positively and was broadly considered to be a thorough process for protecting public sector data. The use of Safe Havens or TREs for storing and accessing such data for research was also reassuring, however there was some initial uncertainty about how they operate in practice and the differences between regional and national Safe Havens. There were also questions about what happens to the data once accessed. While it was clarified that that public sector data cannot be taken out of a Safe Haven , concerns remained about organisations being able to share it with or sell to other companies or countries with looser data protection legislation than the UK.

> "But there's still a lot of unscrupulous people getting it [data]... because I get scam phone calls all the time and that's somebody buying my number." (Participant, Workshop 1)

While some felt that the presentations addressed many of their initial questions and concerns about access to public sector data for research, concerns around data security persisted throughout the workshops.

Presentation 2: data ethics and public good

A representative from RDS gave a brief overview of data ethics and covered the key aspects of defining public good in relation to data research. Public good was defined as work that can improve society by providing a clearer picture of what is needed, such as informing the public about social and economic matters and assisting in policy development and evaluation.

Participants found the focus on using public sector data to deliver tangible societal benefits both interesting and important. Some were surprised that public good in this context is not a widely known or discussed concept. There was broad agreement that research should aim to positively impact people's lives and positive reactions to the idea that researchers must clearly state their commitment to public good before being granted access to such data.

> "I think it's quite positive to see that [researchers] are actually taking information and being precise and transparent, because I don't think that's promoted very well. Their being very precise in that sense is quite refreshing." (Participant, Workshop 1)

However, questions were raised about how exactly public good should be defined and measured in practice. There was a perspective that what constitutes public good is subjective and potentially influenced by political ideology, which prompted a sense of scepticism that criteria for public good could be defined. Concerns were also raised about the potential for organisations to gain access to public sector data based on a commitment to benefitting certain groups but to later renege on that commitment and use the data to marginalise and harm vulnerable demographics. This became a strong theme in later discussions. Questions also arose on whether research could be in the public good if it only benefits a minority of the public.

> "Is there a kind of cut off – does it have to benefit X amount of the public before it's determined to be in the public good? Or is any amount of good, good enough?" (Participant, Workshop 1)

Presentation 3: RDS, Researcher Access Service and Private Sector Access Framework

In the final presentation, participants were introduced to RDS and its role. Delivered by a representative from RDS, it provided a brief overview of the RAS and PIAG as well as plans for developing a Private Sector Access Framework.

Overall, participants were positive about the role of RDS and the use of public sector data for research. RDS's mission to simplify data access for researchers while ensuring projects deliver public good was well-received. Participants saw value in facilitating the process of accessing data for public sector researchers and in having a robust application process. However, there was an appetite to understand more about how applications are evaluated and approved and how RDS gets access to the data. RDS's status as an independent government-funded charity was also not clear to all, and prompted some questions about the role and potential influence of the government in application decisions. These questions were clarified for participants in the later sessions through more detailed presentations and Q&A sessions with representatives from RDS.

> "I suppose it all seems great that small organisations may get access to this data but what about organisations that don't get access? What would be the reasons for this apart from the technical reasons? Are there political reasons? Does government set its own agenda in deciding who can and can't get access to data?" (Participant, Workshop 1)

Reflecting on the initial information provided about the Private Sector Access Framework, there was some scepticism around how commercial interests could align with public interests. This prompted questions about RDS's pricing model and how this would apply to private sector organisations.

> "I just think businesses are making money, so how does that fit in with the public good that they're meant to be thinking of?" (Participant, Workshop 1)

At the end of the first workshop, participants were asked to complete the sentence **"To me, using data for research is in the public good if…"** and their responses are summarised below. Some of these themes (such as not harming any individuals or groups and ensuring inclusivity and representativeness) were further emphasised and built upon as deliberations progressed. Others, such as ensuring tangible and measurable benefits, became more nuanced. Themes around data privacy and security, while remaining important, were less of a focus as participants became more reassured around the safeguards in place. However, some concerns remained and there was a lingering doubt over the efficacy of such safeguards (with recurring references made to data breaches and cyberattacks reported in the media). These concerns were reflected in the views shared on the use of public sector for research in the public good.

Participants' initial reflections on using data for research in the public good

TO ME, USING DATA FOR RESEARCH IS IN THE PUBLIC GOOD IF...

IF IT'S SEEN TO BE INDEPENDENT OF GOVERNMENT

IT IS COST EFFECTIVE

IF IT IMPROVES THE LIVES OF THE PEOPLE IN SCOTLAND

IT IS NOT BIASED

IF THE RESULTS ARE USED TO MAKE EFFECTIVE CHANGE

IT IS ETHICAL

IT HELPS US UNDERSTAND OUR PEOPLE BETTER

IT IS SECURE

IT IS INCLUSIVE, NOBODY IS MARGINALISED, NOBODY IS LEFT OUT

IT TARGETS RESOURCES AND SERVICES BETTER ALIGNED WITH WHAT PEOPLE NEED

IT IS NOT OPEN TO MISINTERPRETATION

IT'S HANDLED RESPONSIBLY

IT MAKES GOOD USE OF LIMITED RESOURCES

IF IT PRODUCES RESULTS THAT CAN HELP PEOPLE

IT'S SECURE

IF IT PRODUCES MEASURABLE BENEFITS

IF IT DOESN'T CAUSE HARM TO OTHER PEOPLE

IT TRULY BENEFITS THE PEOPLE OF SCOTLAND AND WIDER

GIVEN FOR CLEAR REASONS AND ACCESS CAN BE REVOKED IF NEEDED

IT IS REAL WORLD TANGIBLE, SUSTAINABLE OUTCOMES

ACCESS IS UNDER STRICT PROTOCOLS

IT PRODUCES MUTUAL BENEFITS FOR BOTH PARTIES (RESEARCHER AND INDIVIDUAL)

IT IS GOING TO HELP PUBLIC SERVICES IN THE FUTURE

IT IS FACTUALLY CORRECT

IT DEMONSTRATES WHAT ITS OUTCOME IS

IF IT HAS THE PUBLIC INTEREST BEST AT HEART

TRANSPARENCY AND SIMPLE LANGUAGE

IT ENCOURAGES MORE FOCUSED RESEARCH TO BE UNDERTAKEN

IT CAN BE SUBSTANTIVELY MEASURED

IT IS DONE IN A SAFE AND SECURE MANNER

IT IMPROVES THE HEALTH AND WELFARE OF ALL AGES OF SOCIETY

IT IS OF GOOD AND ACCURATE QUALITY IN THE FIRST PLACE

IT ISN'T SHARED FOR COMMERCIAL GAIN

THE DATA IS ACCURATE IN THE FIRST PLACE

THE RESEARCH IS REPRESENTATIVE OF SOCIETY

IF IT CAN BE GUARANTEED THAT THE LONG TERM USE OF DATA IS ETHICAL AND MONITORED

THE RESULTS ARE CLEARLY AVAILABLE TO THE PUBLIC SO THEY CAN SEE WHAT THE BENEFITS ARE

IT IS GATHERED ETHICALLY

RDS AND ITS ACTIVITIES ARE BETTER KNOWN TO THE PUBLIC - PUBLIC MIGHT FEEL MORE CONFIDENT IN THEIR INFO BEING SHARED

ALL RELEVANT DATA IS BEING CONSIDERED

THE PUBLIC IS HAPPY AS A RESULT

IF IT'S FOR THE BETTERMENT OR SUPPORT OF MARGINALISED COMMUNITIES

POLITICALLY NEUTRAL

THE DATA HANDLING IS TRULY ETHICAL

THE PUBLIC BENEFITS AS A RESULT

After the first learning session, participants expressed interest in knowing more about who makes the decisions about granting access to data and how, the specific criteria currently used to define public good and RDS's plans for accessing private sector applications.

The remaining chapters summarise participants' detailed deliberation under each research objective, and the conclusions they formed.

Defining public good in accessing data for research

This chapter summarises how participants engaged with the concept of public good. It presents their conclusions on how public good should be defined when public sector data is accessed for different kinds of research (not just health) and by different kinds of organisations.

Information presented in the workshops

To inform discussions, participants heard two presentations:

- In workshop 1: a brief overview of data ethics and introduced participants to the general concept of public good in relation to research.
- In workshop 2: a more detailed presentation on the RAS, including the process of how data request applications are currently being assessed, with a focus on public good criteria, who is involved in decision making and RDS's future plans to expand its services.

There was widespread support for researchers being required to explain how their research would be in the public good. Many felt that public good was a useful concept, although also recognised its ambivalence and subjectivity making it hard to define. There was a broad view that more clarity and transparency around public good criteria is needed.

> "I think that that for me, [public good] needs to get clarified. Because it's quite a broad term, isn't it? It's quite nonspecific and I suppose, for different topics it will be measured differently. But just in general, how do they [RDS] even gather that kind of information?" (Participant, Workshop 1)

Exploring the concept of public good through scenarios

To help participants explore the concept of public good, they were presented with six hypothetical scenarios involving data access requests from different types of organisations from both the public and private sectors. Participants were split into smaller breakout groups, with each group reviewing two of the six scenarios and each scenario being reviewed twice overall.

For this exercise, participants were asked to take on the role of panel members reviewing a hypothetical research access request. Based on the information provided about the research, they were asked to think about whether or not the research would be in the public good, through weighing up the benefits and risks, and what further information they might need to help decide this. For every scenario, participants were asked to assume that the project had met the Five Safes requirements, particularly in terms of safe project, people, settings, data and outputs and that the data was deidentified but linked at an individual level. They were also reminded that each scenario was made up for the purpose of this research, but aimed to reflect what a real research project might involve.

An overview of each hypothetical scenario is summarised below (full scenarios as presented to participants can be found in Appendix E):

- Scenario A - a university researcher requesting access to social services records, mental health service usage, and data from schools to understand experiences of children in kinship care.
- Scenario B – a public body requesting access to youth and adult justice records, education records collected by schools, and employment status data to understand how childhood interactions with the justice system affect people later in life.
- Scenario C – a charity seeking access to household income and benefits data, area and household characteristics, and education data collected by schools to explore links between income, benefits, and education outcomes.
- Scenario D – a private tutoring company requesting access to education records, household income and benefits data, and household and area characteristics to explore how factors, such as family income, relate to learning challenges and educational disadvantage.
- Scenario E – an international housing development company seeking access to local development and land availability data, and household characteristics linked with income data to identify areas with the greatest housing needs.
- Scenario F – a private energy advice provider requesting access to property and land ownership data, and household and area characteristics to identify properties most in need of energy efficiency improvements.

The following sections summarise the key themes that emerged across these scenario-based discussions, and which formed the basis of participants' conclusions on defining research for public good.[5] Points raised in relation to a particular scenario are highlighted where appropriate.

Benefits of accessing public sector data

Participants recognised multiple benefits associated with allowing access to public sector data for research purposes. These ranged from improving specific interventions, to broader societal impacts.

---

[5] Scenarios D-F were also used to support discussions around models for private sector access to data (summarised in a later chapter).

Targeted interventions

A key benefit highlighted by participants was the potential for research using public sector data to lead to the development of targeted interventions and more effective support for disadvantaged groups. This was seen as particularly beneficial for widening access to services by identifying people with specific needs, such as students with anxiety, children in kinship care, or households experiencing fuel poverty. Research that was seen to lead to more investment, resources and targeted policy action for low-income households and children from disadvantaged backgrounds were also viewed positively.

Service and resource provision

The potential for research insights to be used to direct resources and services to areas of greatest need was another common theme. Access to criminal records, for example, were expected to identify and encourage targeted support to areas with the highest crime rates. This was seen as potentially beneficial for promoting safer communities and reducing the societal and economic costs associated with crime. The potential for research to identify cost-saving measures and support more efficient delivery of public services was also seen as a significant benefit. Participants felt that research providing a more nuanced understanding of the factors influencing social issues could enable policymakers and service providers to allocate resources more effectively and efficiently.

> "I think financially as well, for the country as a whole, it's very expensive, the system of dealing with offenders. So if you can cut down the number of people offending and repeat offending, you're actually benefiting the population as a whole financially." (Participant, Scenario B, Workshop 2)

Participants highlighted the potential for research to drive evidence-based improvements to public services and policies more widely. For example, research requiring access to household and property characteristics to identify properties in need of energy efficiency improvements was seen as potentially beneficial for advancing understanding of fuel poverty and informing energy policy to address it. Similarly, accessing education data was seen as having the potential to identify gaps in education provision, improve understanding of educational inequalities and encourage government action to address them. This was seen as potentially leading to positive impacts and better life outcomes for those perceived to have been left behind or falling through the cracks of the education system. These potential outcomes were seen to lead to wider benefits, such as higher qualification attainment and employability, resulting in economic growth.

> “Well, everybody benefits because education's got wide externalities and it's not just good for the people involved, it's good for your communities, it's good for Scotland, it's good for society as a whole if you can get people out of that deprivation cycle.” (Participant, Scenario C, Workshop 2)

Wider social, environmental and economic benefits

Participants recognised the potential of research using public sector data to lead to wider social, environmental and economic benefits for Scotland as a whole. Research offering more wide-reaching intended outcomes were viewed more positively. The data request of EcoEnergy (a fictional energy advice service) was a particular example of this. As well as helping identify those would benefit most from energy efficiency improvements and making financial support more accessible, participants acknowledged a potential broader impact of the project on the environment and contributing to Scotland's net zero targets.

Participants highlighted the value of research in improving public understanding of complex social issues and informing public debate. It was generally felt that research had an important role to play in addressing knowledge gaps, providing valuable new insights and understanding lived experiences. This was seen as crucial for challenging assumptions, dispelling myths, and building public support for evidence-based solutions.

Risks around accessing public sector data

While there were multiple benefits associated with providing access to public sector, participants also highlighted potential risks.

Potential for stigma or harm

The potential for research findings to lead to the stigmatisation, misrepresentation or stereotyping of certain groups and communities was commonly highlighted. This risk was seen by some as particularly high for groups perceived to be already marginalised, such as ethnic minorities, immigrants, offenders, and young refugees. When considering the use of criminal justice and children's data, for example, there was a perception that research findings could reinforce negative perceptions about certain areas or groups. It was also suggested that such research could lead to further unintended consequences that might harm the communities in question, such as increased surveillance and policing.

> "There could be a risk towards ethnic minorities, or immigrants, or refugee young people. So, there will be a stigmatisation there and it could cause a heightened surveillance by the public and also by the police of these minorities." (Participant, Scenario B, Workshop 2)

The potential for research to be used in ways that harm groups was also identified as a potential risk. Considering the potential for access to public sector data around people's income or receipt of benefits for research, concerns were raised that this could have unintended consequences such as benefits being withdrawn or support for low-income families being reduced. These were perceived as negative and harmful outcomes, reinforcing existing inequalities rather than reducing them.

Misuse of data

Participants were not always confident that the organisations requesting data would have the right level of expertise to draw fully informed and unbiased conclusions from data, particularly where the data or the issue being explored were considered sensitive or complex (for instance, data on children's mental health and social services records). Participants expressed concerns over data requests seeming too narrow in scope for projects focused on exploring the drivers of complex social issues, particularly in relation to children, potentially leading to oversimplified, distorted or incomplete conclusions. This was a particularly strong concern among those who discussed the hypothetical scenario of a charity exploring the links between income, benefits and education outcomes. There was some scepticism around the ability of the research to sufficiently capture the complexities and contributing factors that can affect school attendance (such as medical circumstances) based on the data requested. It was felt that this could result in an incomplete picture of the issue. To mitigate this risk, participants suggested that the panel reviewing the request could issue guidance to broaden the data scope, to consider wider context, or to bring in the perspectives of communities and groups being researched as appropriate.

Concerns were also raised about the potential for researchers to cherry-pick or misrepresent data to fit desired conclusions or advance commercial interests rather than presenting objective findings. A key risk identified by participants was public sector data being used for purposes beyond what was originally set out in the research objectives, particularly in relation to data requests made by private sector researchers. Reflecting a general distrust in private sector organisations, there was a concern that project aims could be intentionally vague to enable the use of data beyond its original purpose.

Despite the specialist presentations detailing the safeguards in place to protect public sector data, concerns around data privacy and misuse were still raised. In particular, participants questioned the capacity of private sector researchers to handle data responsibly. Although they heard from specialists that data cannot be removed from the Trusted Research Environment, participants remained concerned about the perceived risks of unauthorised access or data being shared with third parties.

> "I think a risk is that they [the organisation] would give it to a third party. They would give it to a consultancy to say, 'I've got this. Don't know what to do with it but I think there's something in it here.' It could be shared further." (Participant, Scenario D, Workshop 2)

Commercial gain

Participants considered any use of public sector data primarily for commercial gain, rather than for the benefit of the public, to be unacceptable. Examples included marketing or identifying and targeting areas and groups that would lead to higher profits. This risk was seen as particularly high for data considered to be sensitive (such as income, benefit or social services data), or projects with a focus on vulnerable demographics whom participants felt could be exploited for profit. It was highlighted as a particular concern among those reviewing a hypothetical scenario of an international development company looking to access public sector data to identify areas of land with the greatest housing needs. While addressing housing shortages was one of the stated objectives of project, participants questioned whether this was a genuine aim or a "spin" to disguise the true objective of maximising profits. The underlying concern was that the company would prioritise building expensive housing in high-value areas while neglecting the needs of low-income households, thus exacerbating existing inequalities.

> "It feels like it's to give them an edge over the competition or something like that, you know? To give them a leading edge so that they can target the market or enhance public image – 'we're doing this for the community'" (Participant, Scenario E, Workshop 2)

Weighing up benefits and risks to determine public good

Through weighing up the benefits and risks of different hypothetical uses of public sector data for research, participants began to consider when the use of deidentified public sector data for research was in the public good or not. A central theme throughout the deliberation was the importance of balancing benefits and risks. Participants generally felt that for a project to be in the public good, the expected benefits should clearly outweigh any risks or potentially negative consequences (see Appendix F for an overview of how participants at the in-person workshop weighed up the risks and benefits of the scenarios they reviewed).

Projects that were deemed to be in the public good were associated with clear societal benefits, such as improving health outcomes, addressing social inequalities, improving public services or supporting economic growth. Research that targeted support to groups that are disadvantaged or those with the greatest needs as well as projects that generated wider benefits for Scotland as a whole were seen as particularly valuable and likely to be in the public good.

On the other hand, projects with a perceived high risk of stigmatising or exploiting vulnerable populations were generally not considered to be for public good, raising concerns that the data might be ultimately used to harm or disadvantage the groups being researched. Participants felt that to confidently determine whether the project was truly in the public good, they would need assurances from the organisation that the research would not result in groups being harmed or stigmatised. It was suggested that removing information that could potentially lead to stigmatisation (e.g. ethnicity) from the data request, could help to mitigate these risks.

> "The risks are quite significant. One of the things I'd probably say is [ethnicity] data would have to be removed from the research for this one to stop the stigmatisation. But because a majority of people are affected by crime at one point or another in their lives, the benefits far outweigh the risks of that [project] to me." (Participant, Scenario B, Workshop 2)

The need for clear pathways from research insights to practical action was another key theme highlighted in relation to public good. Participants noted that the value of research ultimately lies in its ability to make a positive difference to people's lives, and that this requires resource availability and effective mechanisms for translating findings into policy and tangible action. This was seen as particularly important for projects targeting vulnerable groups. Participants stressed that the benefits of the research would be limited if the findings were not used to drive meaningful change.

> "I think that there's a lot of benefits there and if it was done properly after the research and something was done about giving everyone support that needed it, then I think it could benefit everyone greatly."
> (Participant, Scenario A, Workshop 2)

There was also a dominant perception that for research using public sector data to be deemed in public good, the benefits should be made accessible to the intended beneficiaries and should not exacerbate existing inequalities. There were concerns that certain groups, particularly those already facing disadvantage, could be left behind if research findings are not used to drive inclusive policies and interventions. Some also felt that the intended benefits should be sustainable and long-term, with research informing systemic changes and improving policies and services perceived as most valuable.

Although initial discussions focused on the idea that research should lead to a tangible, direct benefit, this view became more nuanced as participants reached their conclusions. There was a broad view that research could still be in the public good even if does not directly lead to change or action, as long as a clear roadmap was defined from the outset to show how the insights resulting from the research could ultimately lead to public good being realised in the future. Another view was that while research can be exploratory, it should still outline intended public benefits even if they are not immediate. This view was grounded in the concern about organisations making research objectives intentionally vague to try and conceal the real motives. Ultimately, it was felt that that research in the public good should add value. Overall, indirect benefits were felt to be acceptable as long as a commitment to public good and transparency around the research objectives and intended outcomes at each stage were maintained.

> “If it's like a layered research where the first research is like an indirect one and then it can lead to further research to be done, then it should be clearly explained in the roadmap of the research objectives."
> (Participant, Workshop 3)

In general, participants expressed greater support for public sector organisations and universities requesting access to data. This was based on an inherent expectation that such organisations would be driven by a public good motive. Research proposed by private companies was met with more scepticism as profit-driven motives were often seen to be the primary driver. Where a private company had a more convincing public good image, such as social enterprises, participants tended to view it more positively.

Wider considerations

There were a number of wider considerations that participants felt were important for ensuring research using public sector data is in the public good.

There was a clear appetite for transparency and clarity about research objectives from the outset. It was strongly felt that researchers must be upfront about their goals and justification about why they need the data, what data they need and how they will use it, with commitment to public good being evident throughout. This was seen as essential for ensuring public trust is maintained and to allow scrutiny of the process and outcomes.

> "[As RDS] I'd want to have confidence that whatever the objectives were at the start of the project they've actually been implemented. Having clear objectives and making sure [they] actually get achieved [and] the motives are right." (Participant, Workshop 2)

Participants also sought reassurance that appropriate safeguards would be put in place to mitigate against the risks they had identified (such as harming or exploiting groups, misuse of data or changing scope, or prioritising commercial interests).

A clear understanding of the organisation requesting the data was another safeguard that participants felt would be important. When discussing the scenarios involving a private sector organisation, participants strongly felt that they needed to know more about the organisation to determine if the project was truly in the public good, including things like its environmental, diversity and inclusion practices, values and track record of community investment and contributions to the local economy.

The importance of research being publicly available and accessible was also emphasised as key consideration, particularly when undertaken by private sector companies (see private sector access chapter for further discussion). Details on how and when the findings would be disseminated, and to whom were also deemed important for ensuring that research is being done for public good. Participants called for findings to be openly published and communicated in ways that are meaningful and engaging for affected communities and the wider public. This was seen as important for ensuring transparency, enabling scrutiny, facilitating informed public dialogue and promoting greater public understanding of the issues at hand. Participants also generally wanted to see some accountability and a requirement for the researchers to report back to RDS on their progress against the original research aims and key performance indicators. However, the challenges and limited feasibility of such monitoring were acknowledged, particularly where impacts were expected to be longer term rather than immediate.

Participants' conclusions

At the end of the final session, participants revisited the overarching question of **"How can we tell when the use of deidentified data about people living in Scotland for research is in the 'public good', and when it is not?"**

Their conclusions are summarised below, with aspects of the conclusions that were particularly important to participants denoted with an asterisk (*).

Use of data is in the public good when:

***The research can demonstrate that it will result in a better understanding of an issue by addressing a knowledge gap or providing valuable new insights or demonstrating a specific and measurable benefit for the public.** Some examples might be:

- More affordable housing in the right places
- A fairer and more supportive education system
- Support programmes to promote positive life outcomes for young offenders and help break the cycle of reoffending
- Contributing to Scotland's effort to reach net zero
- Better understanding of the needs of vulnerable groups, like young offenders or children in kinship care

The intended benefits are realistic, actionable and implementable. Where the research is exploratory and is intended to be a stepping stone to future public good, there is a clear roadmap where longer-term potential for public good realisation and intended public benefit are defined, even if not leading directly from the research, for example:

- Identifying areas where affordable homes are needed, and building them
- Improving mental health services for young people
- Targeted interventions in schools for children and young people who need it

***The research prioritises the wellbeing of people in Scotland, and the financial interests of companies are a minor consideration.** For example through:

- Reducing energy bills for households
- Improving the condition of homes (and subsequently the health those living in them)
- Improving support to children and young people in education
- Housing developments that boost local economies and provide community benefits

It is clear who benefits, whether it is a small group or a broad range of people, and the benefits are accessible to those groups.

***The risks of the research are proportional to and outweighed by the potential benefits, and appropriate safeguards are in place to protect vulnerable populations.**

The research maintains public trust and confidence, with data being used responsibly and ethically and original research objectives are adhered to.

Where measurable benefits are expected, timescales for realising these are defined.

Use of data is not in the public good when:

The research objectives are intentionally vague, or the intended public good is unclear and hard to measure.

***There is a risk of doing more harm than good,** for example:

- Singling out, marginalising or stigmatising certain groups, such as low-income households, vulnerable people, or young offenders
- Increasing energy bills of vulnerable households

***There is potential to target vulnerable people or "make money from suffering",** for example:

- Targeting financial products at vulnerable households
- Reducing or removing benefits that people rely on

The intended outcomes OR the roadmap to potential future benefits seem unrealistic or not likely to work in practice. For example, advertising services or products (like installing a more energy efficient heating system) that cannot be delivered due to a lack of supply chain or infrastructure

Profit seems to be the main motivation, rather than a genuine desire to benefit the public, for example:

- Housing developments that are not affordable or do not address the needs of the community
- Higher profits for energy companies with no financial benefits to customers

The research does not lead to the creation of new knowledge or a better understanding of an issue.

Reflecting earlier discussions and considerations that participants highlighted as important, the conclusions were underpinned by the following assumptions:

- The data requested is sufficient to address the research question properly, considering a wide range of factors and perspectives. At the same time, the data requested is not excessive or unnecessary to achieve the stated aims.

- There are checks in place to mitigate risks, and monitor for harms, and address any problems that arise. There is monitoring against misuse of data and results.
- The motivations behind the research, the research process and expected outcomes should be transparent, open and accessible to the public.
- Any significant financial interests are disclosed (especially for private sector research), including funding sources and potential conflicts of interest.

Public involvement in decisions about access to public sector data

This chapter summarises participants' views on public involvement in decisions about access to deidentified public sector data. Specifically, it covers the PIAG being piloted by RDS and participants' conclusions on how this should work.

Information presented in the workshops

To inform discussions, participants heard two presentations:

- In workshop 2, a presentation by RDS gave a brief introduction to the PIAG to set the scene for later discussions.
- In workshop 3, a presentation by RDS introduced further detail on the PIAG, including practicalities (e.g. recruitment, training, compensation), what information about projects would be shared with the Group and when the Group will review applications. The presentation noted that given the volume of applications, not all projects would be reviewed by the PIAG.

After each presentation, participants had the chance to reflect and discuss in small groups and ask questions. Participants were generally very positive about the idea of involving members of the public in the RAS, whilst raising questions about what had prompted its creation and the practicalities of establishing and running the group. There was a strong initial emphasis on ensuring diversity and representativeness in group membership which carried through later discussions.

Benefits and drawbacks of public involvement

Two key benefits of the PIAG were identified by participants: increased public trust and confidence in the process, and improved quality of decision-making. The PIAG was seen as an extra layer of independent scrutiny which could reassure the public that data is only being accessed and used in a responsible way.

> "It will give the public an element of confidence in terms of the process that there is an independent body that will in essence monitor data access with the organisation. From a public perspective that's quite important to have out there" (Participant, Workshop 3)

Participants expected that group members would have a diverse range of perspectives. They felt that this would help challenge potential biases in the decision-making process and ensure that research was in the public good. One participant noted that a public perspective can provide a "common sense check" in addition to the more academic viewpoints of researchers. This was particularly noted in reference to the youth justice system engagement scenario where participants felt it would be beneficial to consult members of the public who had experienced the issues raised in the scenario.

> "You need to also make sure that the people you're asking, or certainly a good percentage of the people you're asking, have experienced some of the issues in those areas as well. If you ask somebody who lives an idyllic life in the middle of nowhere, they're probably going come back with very different thoughts and views on that than somebody who's actually experienced it firsthand." (Participant, Workshop 2)

While participants highlighted benefits of the PIAG and public involvement, they also acknowledged some drawbacks and challenges. This included potential for delays to the research process, and concerns about the lack of expertise or potential for bias among members. Participants noted that members may not be able to fully understand data access requests particularly where these are very technical. The possibility of private sector organisations being able to request access to deidentified public sector data also prompted participants to raise the possibility of members of the public having a negative bias towards these requests (reflecting their own initial hesitancy about private sector access, as discussed in the next chapter).

> "If it requires specialist input that me as a general member of the public might not have, would my input be helpful in a situation like that? There might be instances where you need medical or expert involvement rather than the general public." (Participant, Workshop 3)

> "I think we need to be careful that it doesn't become a kind of bog where things get stuck and we go around the houses and know it just slows things down for perhaps no good reason." (Participant, Workshop 3)

Despite these potential challenges there was an overall sense that the PIAG would be a positive addition to the RAS.

Recruitment and membership of the PIAG

To mitigate potential bias and ensure different perspectives are considered, there was a strong emphasis on the importance of the PIAG representing a broad cross-section of society.

> "An elderly pensioner's opinion may differ very much from a younger person or a young family. You need to make sure you've got that full range of age groups that you're speaking to." (Participant, Workshop 2)

Some participants specifically mentioned the need to include individuals who might not typically engage in such activities. As touched on above, participants also highlighted the value of membership of the PIAG including experts such as specialists in data ethics, research methodologies, or with relevant lived experience. Given the need to include a range of experiences and expertise, it was suggested that the group could comprise a “pool” of members who could be selected to review individual applications based on their eligibility. This could include consideration around their interest in the subject matter, lived experience, expertise, or financial interests.

A robust recruitment process was seen as essential to the success of the PIAG. There was strong agreement about the need to widely advertise opportunities to take part. This could involve actively seeking out individuals from underrepresented groups. Some participants suggested a screening process like the one that had been used to recruit them to take part in the workshops. To further reduce the risk of bias and potential for influence, there was a suggestion that group members could be anonymous to applicants (particularly those from the private sector).

There was agreement that group members should receive financial compensation and training. This would ensure that a broader range of individuals can participate, regardless of their socioeconomic background and existing knowledge. Participants also noted that group discussions should be supported by a skilled and independent facilitator (perhaps from outside RDS) to ensure all voices are heard and discussions stay focused.

Scope of involvement

There was an overall preference that the PIAG should review each application. This was seen as important to maintain transparency in the process.

> "I think it's important that if the research is going to impact the public, even in a positive or negative way, the panel should see it before it goes anywhere and they should pass it through. I don't think something that can impact everyone should not be seen by this group." (Participant, Workshop 3)

If this is not possible, the type of organisation and level of risk were the key factors that participants felt should determine whether a project is reviewed by the PIAG. There was a general preference for the PIAG to review all private sector applications due to concerns about the motives of private sector organisations (prioritising profit over public good). There was also a view that this perception could lead to members of the PIAG holding a biased view of private sector organisations, and it was felt that this would also need to be managed to ensure these requests were reviewed fairly and consistently.

> "When it comes to universities, you're more likely to trust them anyway because they tend to be doing something for good. Whereas when it is a private company and there's profit involved, there could be ulterior motives there and they may not give you all the information straight away." (Participant, Workshop 3)

Projects that were seen as riskier included those requesting data on children, research on a political topic, and those requesting "sensitive data", as described by participants, such as on age, gender, and socio-economic status. Some participants felt that there was a risk that members of the public could be identified from this type of data even if it was deidentified, and would want assurances that these are being handled carefully. There were differing views on health data; while some participants felt this type of data was an increased risk, others felt that RDS would have the experience to review these applications without public involvement. Where the application is very complex or technical, one view was that this may not need to be considered by the group, however another view was that the researchers should explain their application in an accessible way so that it can be scrutinised by members of the public.

No issues were raised with the PIAG reviewing applications before the Approval Panel. However, there were questions about who would decide which applications were reviewed by the PIAG. To improve transparency, there was a preference for the PIAG to have sight of all applications, and to have the ability to request to review an application if it disagrees with the decision that PIAG involvement is not needed.

> "If it's decided that something doesn't need to go to the group, then maybe the advisory group should be told, 'we're doing this, by the way, but we don't think you need to see it' and give them an opportunity to say, 'hold on, I think maybe we should have a look at that'." (Participant, Workshop 3)

There was also a preference for the PIAG to be informed about decisions made by the Approvals Panel and able to challenge whether or not a project is approved.

Participant conclusions

Following their discussions, participants revisited the overarching question of public involvement. Their conclusions are summarised below:

How should the public be involved in decisions about sharing deidentified data about people in Scotland?

- Opportunities to join the PIAG should be widely advertised to ensure fairness and representativeness. Advertising could be done via social media, in newspapers, or in public places. RDS should clearly explain what the group is and what membership will involve.

- Diverse representation is essential. The Group must include people with diverse demographics, socioeconomic backgrounds, lived experiences, and expertise. Members could be asked some screening questions before they are invited to join the Group to ensure there is diverse representation.
- Any potential conflicts of interests should be disclosed by members. Members could be anonymous to researchers and organisations to prevent potential influence by them.
- There must be clear guidelines, training and support for members of the Group so they have the knowledge they need. The content of group discussions should be confidential as members will have access to privileged information.
- Skilled and independent facilitation of the Group will be necessary to ensure all perspectives are considered in group discussions.
- Group members should be compensated financially for their time to ensure a cross-section of society can take part.
- Ideally the Group should review all applications. If this is not possible, involvement will depend on the project type and risk. All applications from the private sector should be reviewed by the Group. Even if the Group does not carry out a full review, it must be aware of all applications that are made.
- The Group could be involved at an early stage before proposals are formally submitted as a way to shape research design, identify potential risks, and ensure public good from the outset. The public could also be involved in monitoring to make sure researchers do what they say they will.
- If the RDS Approval Panel reaches a different decision to the conclusions reached by the Group, the reasons for this should be clearly explained to the Group.
- There is a need for greater transparency and public awareness about data sharing and research more generally.

Private sector access to public sector data

RDS does not currently provide private or third sector organisations with access to public sector data, but may do so in the future. Given the complexities around private sector access to public sector individual data, RDS is working with Scottish Government to develop an operational framework that ensures a consistent and trusted process for enabling private sector access.

This chapter summarises participants' views on private sector access to deidentified public sector data. Specifically, it covers views on three possible models for private sector access and the key considerations for each approach.

Information presented and initial views

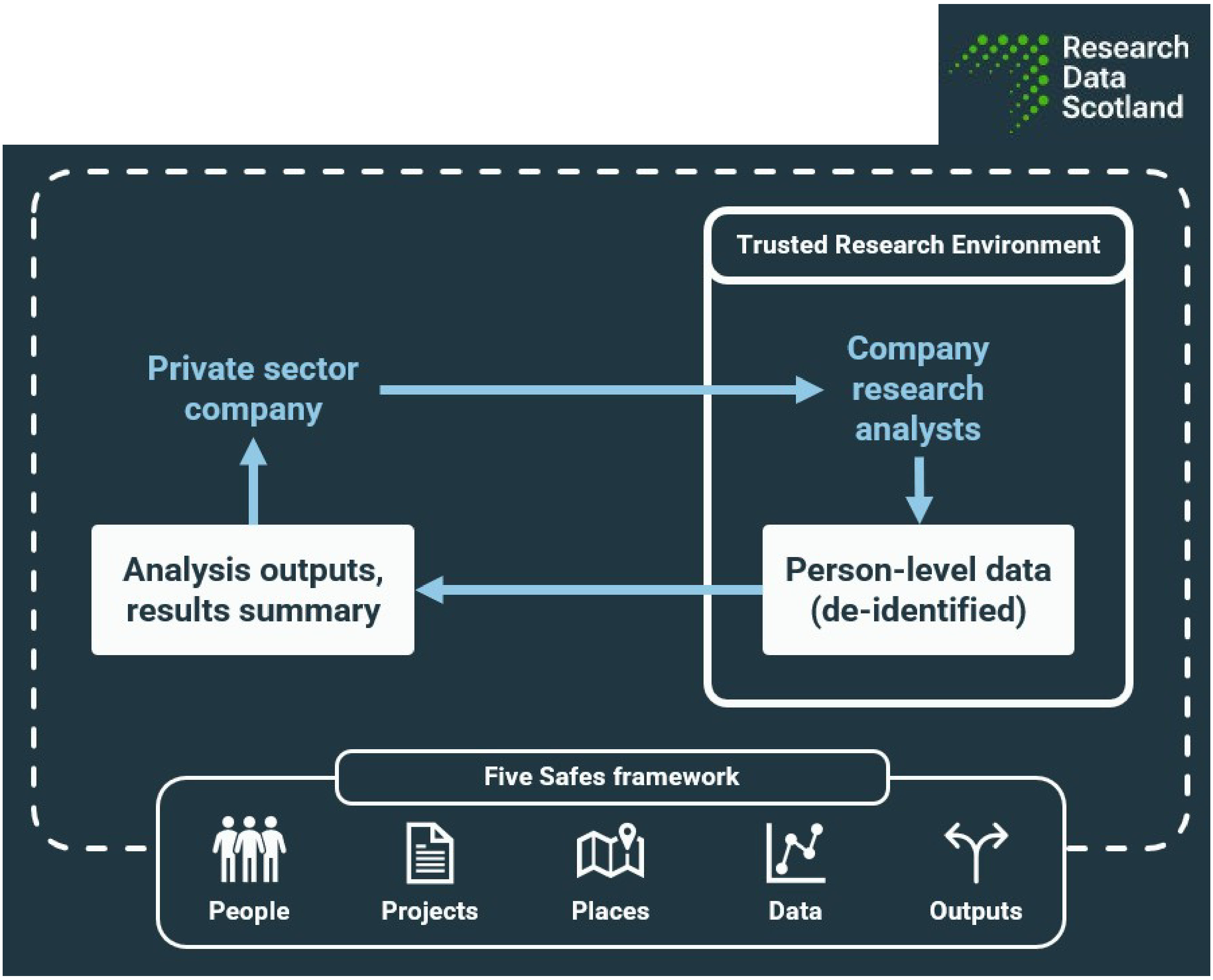
To inform discussions participants heard a presentation in workshop two from a representative at RDS which introduced the concept of private sector access to deidentified public sector data, including the current risks and opportunities of working with the private sector on research involving public sector data. For the purpose of this research, the private sector was defined as any organisation aiming to make a profit.

Participants also heard about RDS's plans for setting up a Private Sector Access Framework to allow private sector access to public sector data in a safe and consistent way. Participants were introduced to three models that could be applied within the Framework:

1. **Model 1: Research by a private sector organisation** where the organisation would access the data directly within a Trusted Research Environment.
2. **Model 2: Research by a private sector organisation working with a public sector organisation or university researcher** where both organisations would access the data together within a Trusted Research Environment.
3. **Model 3: Research for (but not by) a private sector organisation** where a public sector specialist or university researcher would access the data on the private sector organisation's behalf in a Trusted Research Environment.

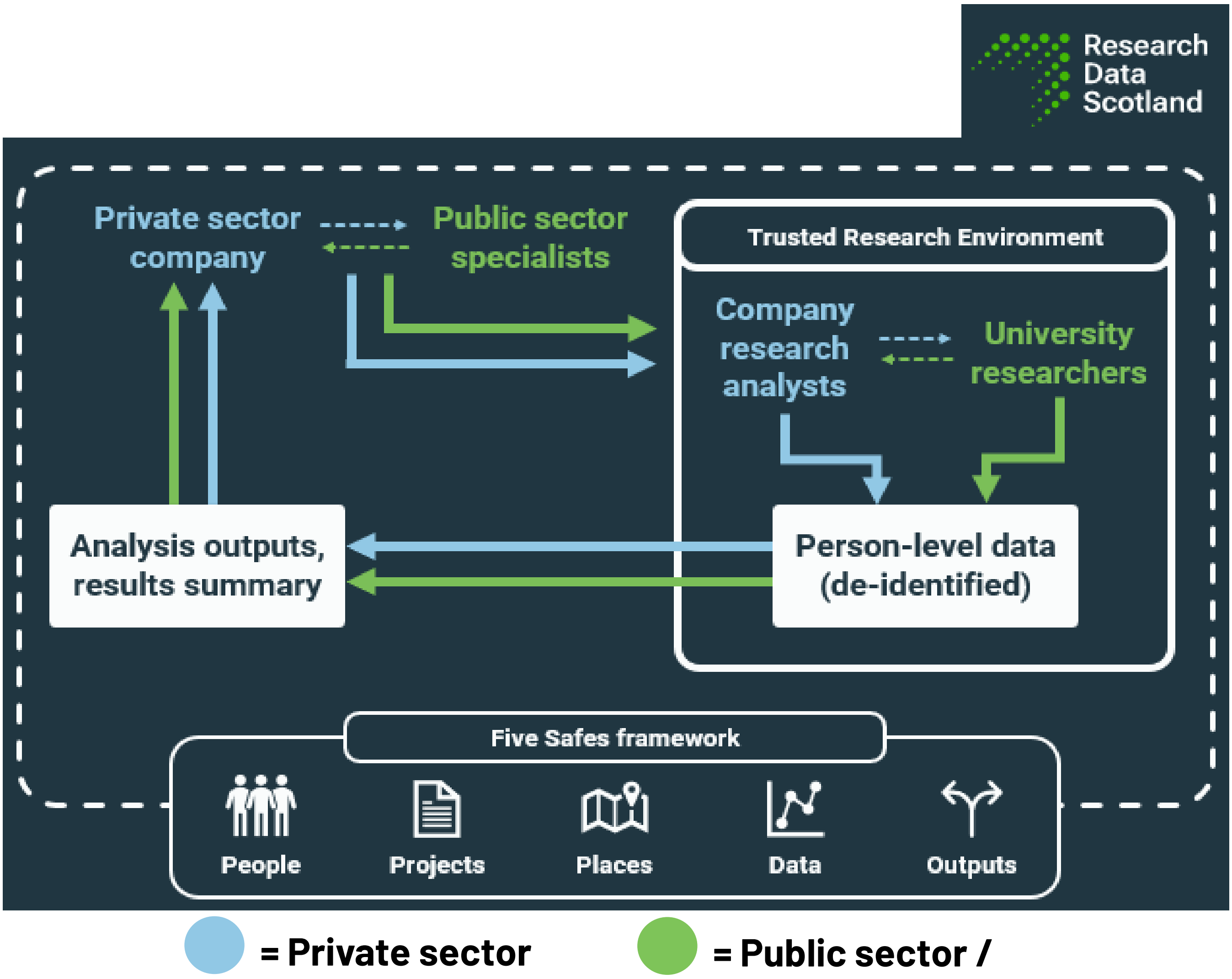
Model 1: Research by a private sector organisation

Researchers from a private sector organisation would access deidentified public sector data directly.



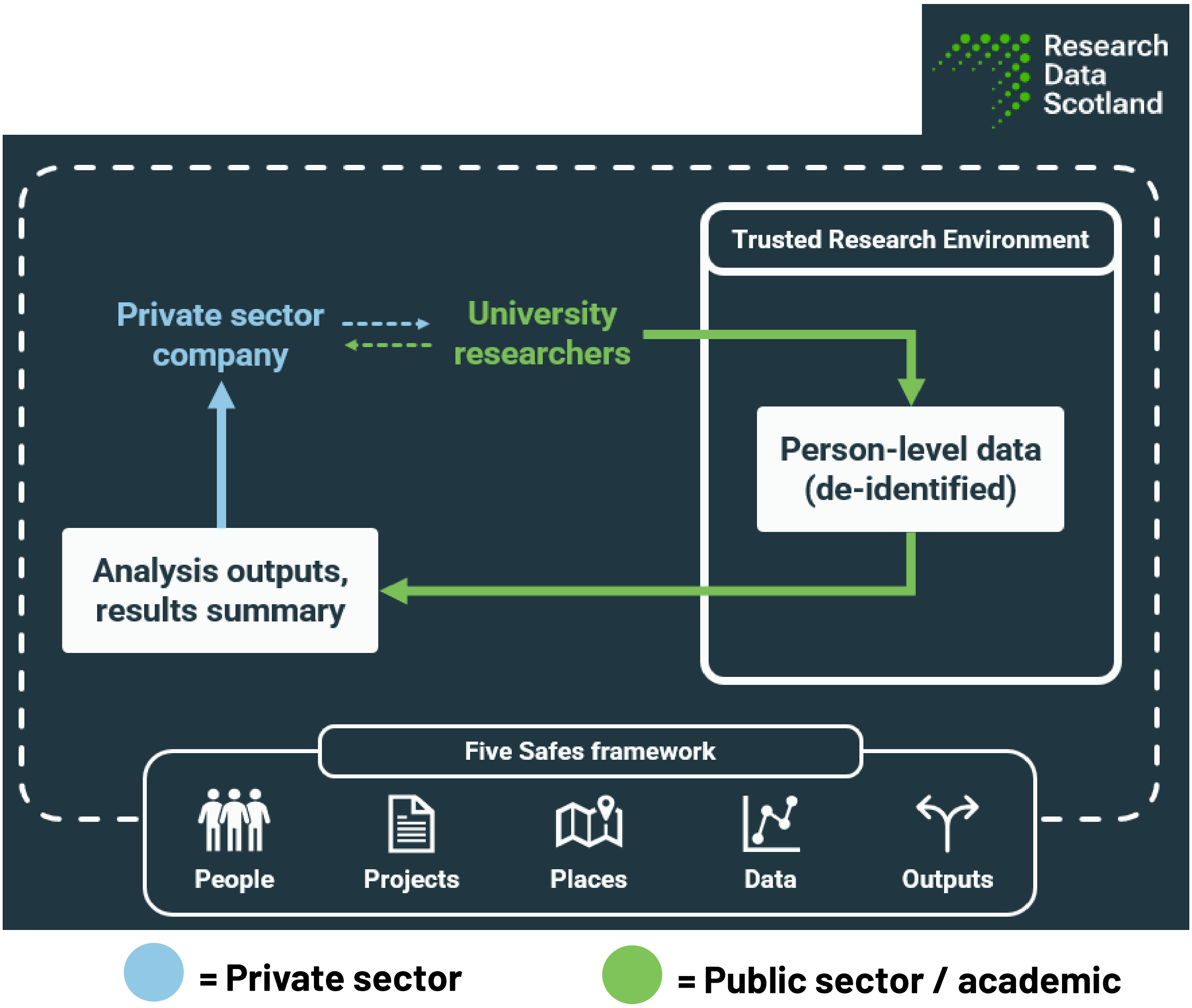
Model 2: Research by a private sector organisation working with public sector and others

Private sector researchers and university researchers work together when accessing deidentified public sector data.



Model 3: Research for (but not by) a private sector organisation

Private sector organisations have the research done for them by university researchers.



After the presentation, participants had the chance to reflect in small groups and ask questions.

A key theme in initial discussions was concern about data security and privacy in relation to private sector access to public sector data. Despite clarifications around public sector data being deidentified, some apprehension remained around the risk of data breaches and leaks to other companies. This was grounded in pre-existing concerns about the sharing of personal information for marketing or scam calls expressed in the early stages of deliberation.

Reflecting earlier discussions around research for public good, participants emphasised the need for private sector research having a clear and demonstrable public good. While there was a general view that the processes for private sector access seemed secure, there were also concerns about private sector organisations having ulterior motives and that this would require heavy scrutiny to determine (see public good chapter). Insurance companies were mentioned by one participant as an example of where use of data could be used to save costs through refusing coverage for people with certain health conditions. At an early stage, participants therefore highlighted the importance of understanding private sector motivation for accessing deidentified public sector data, which became a strong theme over the course of deliberation.

There was broad support initially for the models of access that involved private sector organisations collaborating with universities or public sector bodies. These felt to offer advantages in terms of providing resources for research that could benefit the health and wellbeing of society, whilst also providing robust oversight.

> "I think for me, it's all about value for the public good. That's the peak idea [...] I know a lot of universities do collaborate with a lot of organisations and they do add value in terms of research and R&D in many ways. So, if they make a bit of [profit] off that, I personally don't have an issue with that. So, as long as it's done in the right matter, and in this context, for the public good" (Participant, Workshop 2)

Participants had initial questions around the practicalities of private sector access to deidentified public sector data under these models, including what the criteria would be to determine which model was appropriate, how partnerships with universities would be resourced, how findings from such research would be shared both in Scotland and beyond to improve collective knowledge, and how any profits arising from the research could be shared back with the public.

Following clarification around these points, participants then went on to consider the models in detail.

Views on models for private sector access

To support discussions of each model, participants revisited the hypothetical scenarios that involved private sector organisations:

- Scenario D: a private company called Educ8 seeking access to education, income and benefits, and household data to inform its tutoring services.

- Scenario E: an international housing developer called Nu Homes seeking access to local development and land data to identify local housing needs.

- Scenario F: a private company seeking access to property, land ownership and household data to inform its energy advice service.

Participants discussed the extent to which the risks and benefits they had identified in relation to the scenario (covered in the public good chapter) could be mitigated or ensured through the models, and the key considerations for each one. Although participants were not asked to compare or rank the models, as it was acknowledged that the models had been simplified and would in reality be used flexibly, some preferences were nevertheless expressed and are reflected in the findings.

Model one: research by a private sector organisation where the organisation would access the data directly

Access to data under model one was broadly felt to be the riskiest approach. This view was grounded in a general distrust of the private sector and a perceived lack of scrutiny over what a company would do with the data. There was a general sense that model one would not provide sufficient safeguards against private sector interests. This view was particularly strong among those who engaged with the NuHomes and EcoEnergy scenarios, as it was felt these fictional companies were prioritising profit over public good.

Related to this, a lack of transparency was also highlighted as a risk under this model, with the possibility of private companies “cherry-picking” the findings to support their agendas. Furthermore, risks around data misuse due to shifting scope, inadequate monitoring and lack of capability among private companies to analyse the data accurately were considered to be higher under this model.

However, an alternative view was that private sector companies could have more relevant market knowledge. For example, EcoEnergy was described as an organisation that would have a clearer understanding of the needs of different households through their work. This “boots on the ground” experience was identified as a potential benefit of research by a private organisation with direct access to the data.

Direct access to the data by private sector organisations under this model was also associated with potentially faster public good realisation. This was based on the assumption that private sector organisations would be able to operate more efficiently and would have access to more resources, enabling them to conduct the research and deliver on the public good more quickly than a public sector organisation could.

> "It sits uncomfortably with me, but I can absolutely understand that there just may be resources and capability and skills that would be much better used and afforded privately to process all this, instead of giving it to the public sector/academia." (Participant, Workshop 2)

Although there was cautious acceptance of this model as deliberations progressed, it should be noted that some participants remained wary of it. While acknowledging the benefits, they felt these were outweighed by the risks and so this model was not acceptable for them personally.

Model two: research by a private sector organisation working with a public sector organisation or university researcher

The involvement of a public sector body or university was widely felt to lend credibility to the process. While some participants became more accepting of there being a profit element to private sector access to data alongside a public good, the involvement of a researcher from the public sector or a university was described as a "belt and braces" approach, providing "checks and balances" to ensure the research was primarily driven by a public good purpose and not just commercial interest, such as building affordable housing that addresses local need (as per the Nu Homes scenario).

> "For me personally it gives me a lot more comfort if there's an established public body that was involved. They'll have the interest of the people at heart as opposed to profit. So, they might be able to rein back decisions that [a housing developer] would make if they are profit-driven, they might say, 'well, actually, that doesn't fulfil that need'." (Participant, Workshop 2)

There was also a view that model two offered a more joined-up approach, with the private and public sectors each bringing their own strengths; it was suggested that the rigour of academia coupled with the resource and market knowledge of a company could improve the quality and impact of the research. While it was acknowledged that this approach could be slower and more costly, there was a broad sense that the processes in place within universities would ensure accountability, which participants highlighted as an important part of the process.

> "I think having another body doing the research adds to the level of accountability. If a university was involved, would it have to go through their ethics committee as well as [RDS's] vetting procedures? Because that, again, would make things longer but it would make it more accountable." (Participant, Workshop 2)

Related to this was a view that the private sector partnering with a public sector body or academic would mitigate risks of biased findings and misinterpretation of data. Having previous raised these as potential risks (see public good chapter), participants considered the public sector partner role to be primarily about oversight, ensuring the research findings were objectively presented. This form of collaboration was also felt to mitigate against private sector organisations unilaterally changing research aims or misusing data, which participants continued to express concerns about even with knowledge of the Five Safes Framework. Participants felt the public sector body would play a key role in ensuring the research stayed focused on its stated aims.

However, there were doubts raised over the reliability of public sector bodies, particularly universities, to provide this oversight role. This was grounded in a view that university motives could also be financially driven, and led some to question whether universities could be influenced by private sector funders and lose sight of the project's public good objectives.

> "I can't decide whether it would or wouldn't create more social good because in my experience, a lot of the universities, the financial purpose of them is to commercialise them quickly." (Participant, Workshop 2)

Participants also raised concerns over the ability of private sector and public sector organisations to work together effectively. Conflicting priorities, unclear roles and responsibilities and a lack of trust between organisations were all highlighted as possible reasons why a partnership approach under this model could falter. Underlying this was an inherent distrust of private sector motives and a perception that the private sector would act faster on research findings. In the Nu Homes scenario, for instance, participants were concerned that the housing developer would "outmanoeuvre" its public sector partner and not follow through with its promise of building affordable housing.

> "If two organisations are getting the same information, then the private sector can act on that information a lot quicker than the public sector can. So you say, right, there's 400 acres of wealthy land there. Public body are going to mess about for a year, but the private sector will buy it." (Participant, Workshop 2)

Model three: research for (but not by) a private sector organisation

Participants highlighted similar benefits to this approach as with model two, viewing the involvement of both private and public sectors as a good balance between the innovation and resource of the private sector and the impartiality and rigour of the public sector.

The perceived "separation" between the private sector organisation and the deidentified data, with the data being accessed by a public sector or university researcher, was reassuring to some. It was felt to reduce the risks around misuse of data while still enabling the insights to improve service delivery, such as the tutoring service that was set out in the Educ8 scenario. However, others did not see a significant difference between university researchers or private sector organisations accessing the data, assuming it remained deidentified.

The involvement of the public sector under this model was also felt to have the potential to influence broader government policy decisions. Based on the scenarios considered, participants highlighted the policy areas of energy and land use in particular.

As with model two, participants highlighted potential challenges with the public sector working with the private sector. Of particular concern was the scope for bias and corruption, with questions raised over whether funding from a private sector organisation would compromise the independence of a university. There was also a concern that private sector organisations would “cherry pick” a public partner based on alignment with their motives rather than in the public interest.

> "He who pays the piper calls the tune." (Participant, Workshop 2)

In terms of practicalities, it was felt that this model could be the least efficient due to the additional steps involved for data to be accessed via another party.

Wider considerations

There was a general view that any of the three models would be technically acceptable as long as they met the conditions of the Five Safes Framework.

> "How much difference actually is there though when you look at all the five safes, when you're looking at the trusted research environment, it shouldn't make any difference what researchers are [accessing] it. So it actually makes no difference who's [accessing] it or how many because they've got to fulfil all the five safes." (Participant, Workshop 2)

When discussing the models in relation to the scenarios, some involvement of a public sector body or university was felt to be important for ensuring that the use of deidentified public sector data aligned with public needs ahead of any commercial gain. However, participants’ main considerations around private sector access centred around the purpose of the research and assessing the potential risks and benefits on a case-by-case basis.

> "I think more would need to be looked into the risks first just to see exactly who is going to benefit from this regardless of who actually carries out the research." (Participant, Workshop 2)

Nevertheless, some participants highlighted the importance of scrutinising private sector organisations and having particular safeguards in place through thorough “vetting” or “background checks”. In particular, the company’s ownership structure, reputation, values and previous research were considered to be important for RDS to assess before providing access to public sector data. This was felt to be necessary for protecting RDS’s reputation as well, and it was suggested that the PIAG could have some involvement as a further safeguarding measure.

Participants also emphasised the need for transparency to build public trust in these processes. In particular, it was suggested that there should be a requirement on the private company to publish the findings from any research accessing public sector data. In the final workshop, further clarification was given around what would be published on RDS's Data Use Register which included a summary of the research and lead organisation, a public good statement and details of the datasets accessed. RDS highlighted that private sector organisations may be restricted on how much they could publicly publish due to commercial sensitivity.

Reflecting on this further, participants expressed mixed views on whether the publication of research findings should be required or not when private sector organisations seek access to public sector data. One view was that making this a requirement would give members of the public confidence in the process and trust in the organisations involved. For some it was felt to be important for transparency and a reasonable trade-off for accessing public sector data, even if it was costly to, or off-putting for, some organisations.

> "How can you be transparent if you don't know what they're doing with information? Regardless of whether it suits their business or it benefits them or not, we need to know how they've used the data that they've got. So I don't think it makes any difference whether it's benefited to company or not. We need to know how it's been used, otherwise there's no transparency." (Participant, Workshop 3)

For others, it was felt to be necessary both for ensuring public good realisation and for safeguarding against potential misinterpretation or "cherry picking" of research findings for commercial rather than public interest.

> "If they really wanted to do something in the public good, I don't see why that would put them off. If that puts them off, that's a red flag for what the motive was in the first place." (Participant, Workshop 3)

Another view was that it could be unrealistic to ask companies to make their findings publicly available and it would not be feasible to monitor what companies ultimately discover from the research. Additionally, it was recognised that this requirement could deter some organisations from conducting research that could benefit the public. This prompted wider discussion around the payment model for private sector accessing public sector data and calls for clarity on how this would work in practice.

Overall, there was an appetite to balance the need for transparency with the commercial sensitivities. Participants suggested a number of ways in which this could be done, including:

- Setting a minimum requirement to publish some details, such as an overview of what the research was about and how it fulfilled its original purpose.

- Having a grace period where publication was not required, after which time the research should be made publicly available. The period of time could be agreed on a case-by-case basis.
- Having research findings available upon request, with clear signposting on the RDS website for the public to make such requests.
- Providing information that cannot be published to RDS for review.

Ultimately, while there was an overall preference for research findings to be publicly available, it was felt that this was secondary to research being in the public good and conducted by a trustworthy organisation. It was also suggested that there should be some form of monitoring of the research, with consequences for private sector organisations who do not do what they say they would.

> "It goes back to the benefit of the public good. That's what it has to come down to. And whether the private company wants to make that information public or not shouldn't really be the focus. It should be the public good." (Participant, Workshop 3)

Participants’ conclusions

At the end of the final session, participants revisited the overarching question of **“which methods for private sector organisations accessing data about people in Scotland are acceptable?”**

With the five safes framework in place, it was broadly felt that any of the models would be acceptable, as long as the research is demonstrably for the public good. As outlined above, there were exceptions to this, with some remaining uncomfortable about private sector organisations accessing deidentified public sector data without any public sector or academic involvement.

In summarising their conclusions on each model (presented below), participants highlighted some key considerations for any model of private sector access. These included:

- **Transparency**, to ensure public confidence in the processes used to provide access to deidentified public sector data.
- An understanding of the private sector organisation requesting access, through **thorough vetting** of the company structure, history, values and commitment to public good.
- **Consistency** – what applies to one model should apply to all.

Key considerations for model one (research by a private sector organisation):

- This could be seen as the highest risk model due to concerns about the potential for exploitation of data for profit, lack of impartial oversight, potential for bias in research design and interpretation, and that some private sector organisations might not have the same level of expertise in data interpretation as the public sector.

- But with sufficient safeguards in place, including input from the PIAG, this may be the fastest route to public good realisation if the research is conducted more quickly than the other models, as the public sector might not have the resources to do the same research as efficiently.
- There should be a clear requirement for the private company to make some information about the research publicly available, to avoid the risk of them 'cherry –picking' the findings that align with their commercial motives. At a minimum, this should include information about the purpose and aims of the research, progress towards aims and some findings. RDS should be able to check back against the stated aims to assess whether the company has achieved them or not.

Key considerations for model two (research by a private sector company working with public sector and others):

- Offers a balance between accountability and efficiency.
- Offers checks and balances leading to trust in adherence to the research objectives and public good realisation.
- The collaborative approach has the potential to strengthen the research through combining resources and expertise which can lead to greater public good.
- But there is potential for bias and undue influence with private companies 'cherry-picking' partners who are likely to support their agenda.
- Where a partnership is in place, the organisation funding the research should be disclosed.
- Clear roles and responsibilities of each partner should be defined at the outset.

Key considerations for model three (research for, but not by, a private sector company):

- This may not mitigate risks around private sector access if the private company is funding the research.
- It is potentially the least efficient model resulting in slower research process and public good realisation.
- Safeguards should be put in place to ensure researcher independence and prevent against undue influence from the private sector.
- There should be a clear requirement for the private company to make some information about the research publicly available.

Reflections on key concepts and wider communication

This chapter summarises participant reflections on key concepts and issues discussed throughout the workshops, and shared views on communicating these with the general public more broadly.

After the first and second workshops, participants were sent a short survey to measure understanding of key concepts over the course of the research (see Appendix H). In the final workshop, participants shared their views on the concepts they had found engaging or challenging, and what factors supported their learning.

Concepts of most interest

Reflecting on their experiences taking part in the research in the final workshop, participants highlighted the following concepts that were of particular interest to them:

- **The role of data in society:** the potential for data to be used for research to improve society resonated strongly with some participants.
- **RDS' role and safeguards:** some participants were reassured to learn about the safeguards that are in place to protect data about people living in Scotland that is collected by the public sector. This included the concept of deidentification, Safe Havens, and the checkpoints on a researcher's journey to access data via RDS. All of these were seen as crucial to protecting people's privacy and ensuring public trust and confidence in the system.

Challenging concepts

The following concepts were felt to have been more challenging to understand:

- **Public good:** some participants noted that it was more difficult to understand the concept of public good when it was described in an abstract way. It was easier to understand when discussed in relation to the hypothetical scenarios.
- **The models for private sector access:** some participants felt that the models for private sector access were difficult to understand. It would have helped to have presented examples to illustrate how these models might work in practice when they were first presented.
- **The Five Safes Framework:** some participants recalled that it took them some time to understand the framework, but highlighted the graphic as a helpful aid.
- **Data sources:** there was an appetite for more detail on where the different non-health data sources that RDS is interested in making available would come from.

Communicating key concepts with the general public

As touched on, participants highlighted a range of factors that helped in their understanding of key concepts throughout the workshops. Several participants mentioned that hearing about 'real-world examples' of how data has been used for research, and the resulting benefits, had helped make discussions about concepts more relatable. Real examples of research resonated in particular:

> "If you're not involved in research at all, then you don't know what people ask for and what they do [with it]. I think the explanation about how they identified how women had heart attacks differently to other ones, or the mental health issues of people that were homeless and things like that were useful" (Participant, Workshop 3)

Charts, maps, and diagrams helped simplify complex information and illustrate processes. The map of Safe Havens, illustration of the Five Safes and checkpoints of a researcher's journey for accessing data through the RAS, stood out in particular:

> "On the visual things, I also really liked the map of the Safe Havens. I thought that was really easy to understand because in my head when they were talking about Safe Havens I didn't understand how it was broken up." (Participant, Workshop 3)

More generally, the use of clear and simple language, avoiding jargon, was seen as crucial for making the information accessible to a general public audience. Some participants also mentioned that they had learned a lot of new information during the workshops, so it had been helpful and more manageable to be introduced to the concepts in stages.

Reflecting on what they had learned over the course of the research, participants identified the following key concepts that they felt that RDS should communicate to the general public:

- **The existence of RDS:** many participants were unaware of RDS and its role in providing access to data for research.
- **Public good:** the public could be better informed on how data has been used and how this has contributed to public good. Providing examples of previous projects could help demonstrate the value of research using public sector data.
- **Safeguards:** as discussed, communicating the measures that are in place to protect public sector data and ensure that it is used for public good were seen as crucial to building confidence.
- **The PIAG:** the public should be informed about the existence of the PIAG and its role in ensuring that public views are considered in decisions about data access.

Participants also made the following recommendations for how RDS should communicate these key concepts with members of the public:

- **Use a variety of channels to reach a wide audience:** This could include social media, newspapers, radio, local billboards, and community events, ensuring that information is accessible to different groups.
- **Be open and transparent about processes and safeguards to build trust and confidence:** It was also felt to be important that RDS is open to answering questions from members of the public.
- **Awareness campaigns:** It was suggested that RDS could consider launching educational initiatives, such as in schools or other community venues, to raise awareness of data-related issues and to promote data literacy.

Conclusions

This chapter summarises the overarching themes that emerged from participants' deliberations.

Developing criteria for assessing when access to data is for the public good

There was widespread support for organisations being required to set out how their research would be for the public good before accessing public sector data via the RAS. The challenges in defining public good were acknowledged, given its abstract nature and dependency on the specific research context. Nevertheless, in their conclusions on how to determine when access to data is for the public good, participants identified criteria by which this should be assessed. This was felt to be widely applicable, regardless of the type of data being accessed or type of organisation seeking access and highlighted the following themes:

- **Transparent and tangible benefits** – it was broadly felt that organisations should be able to explain how their research would result in a better understanding of an issue, whether that be through addressing a knowledge gap, providing valuable new insights, or demonstrating a specific and measurable benefit for the public (such as improving health outcomes, addressing social inequalities, improving public services or supporting economic growth).
- **Prioritising wellbeing** – the research should clearly prioritise the wellbeing of people in Scotland, with any financial interests (particularly those of private sector organisations) a secondary consideration. Research that targeted support to groups that are disadvantaged or those with the greatest needs as well as projects that generated wider benefits for Scotland as a whole were seen as particularly in the public good.
- **A considered risk-benefit analysis** – for a project to be in the public good, the risks should be identified and acknowledged, with a rationale for how they are proportional to, and outweighed by, the benefits.
- **Safeguarding vulnerable groups** - it was strongly felt that proper safeguards and assurances must be in place where there is a risk of data being misused or misrepresented to harm, exploit or reinforce negative stereotypes about groups, especially for those already experiencing vulnerability or marginalisation.

Public involvement in the Researcher Access Service

Participants were broadly in favour of the PIAG as a way of involving the public in decisions as part of the RAS. It was seen as an additional layer of independent scrutiny to help increase public trust in the use of public sector data for research. A number of key considerations for the setup of the PIAG arose from participants' conclusions on involving the public, including:

- **Diverse representation** - the PIAG should represent a broad cross-section of society, including individuals from underrepresented groups and with relevant lived experience.

- **Robust recruitment** - opportunities to join the PIAG should be widely advertised, with a clear explanation of the group's purpose and membership requirements.

- **Training and support** - group members should receive financial compensation and training to ensure a broader range of individuals can participate, regardless of their socioeconomic background and existing knowledge.

- **Transparency** – decisions around which applications the PIAG provides input on, and the decisions made by the Approvals Panel on applications, should be made in a consistent way and clearly communicated the Group,

Considerations for how private sector organisations could access public sector data for research

Despite a general distrust of the private sector, there was broad acceptance that access to public sector data by private sector organisations could be in the public good with appropriate safeguards in place. Participants supported some form of partnership working between private sector and public sector organisations to balance resource and innovation with robust oversight, as highlighted in participants' conclusions.

Nevertheless, it was felt that how private sector organisations access the data was secondary to whether or not the proposed research was demonstrably for the public good, and that all researchers should be subject to the same criteria (as above). Particular criteria for the private sector included:

- **Vetting** - a thorough understanding of the private sector organisation requesting access was deemed essential, including its structure, history, values, and commitment to public good.

- **Consistency** - what applies to one model of private sector access should apply to all.

How best to communicate concepts relating to research data projects to the public

Reflecting on their experiences taking part in this deliberative research and the key concepts that they learned about, participants felt that public awareness of RDS and the use of public sector data, safeguarding mechanisms and public good principles should be raised. Their recommendations included:

- **Utilising various channels** to reach a wider audience (such as social media, newspapers, radio) and using real examples to help bring concepts to life.

- **Being open and transparent** about the processes and safeguards to build public trust and confidence. Using clear, accessible and jargon-free language was felt to be important when communicating messages.

- **Awareness raising campaigns** to promote data literacy, through school and community events.

Appendices

A. Overview of workshops

Format	Date/time	Group	Objective	Description	Presentations
Online workshop Introductions and learning	Wednesday 23 April 2025, 6-8.30pm	All participants	Familiarising participants with the process and topic, introducing overarching questions.	Introduction to the process and to those attending. Presentations from subject specialists followed by small breakout discussions and Q&A with specialists in plenary.	Introduction to data for research and the 5 Safes Framework. Data ethics and public good. RDS, RAS, PIAG and private sector access framework.
In-person workshop Scenario discussion and initial conclusion forming	Saturday 26 April 2025, 10am-3pm	Participants living near/able to travel to central Edinburgh location (24)	Explore topics in more detail and deliberate to form initial conclusions on overarching questions	Playback of what's been covered so far, presentations from RDS with more detailed overview of work streams (RAS, public good, PIAG, private sector access framework). Scenario-based discussion to explore public good, public involvement and private sector access. Q&A with specialists in plenary (and at tables).	RAS, public good and PIAG. Private sector access.

Online workshop Scenario discussion and initial conclusion forming	Thursday 1 May 2025, 6-9pm	Participants not living near/able to travel to central Edinburgh location (12)	Explore topics in more detail and deliberate to form initial conclusions on overarching questions	Playback of what's been covered so far, presentations from RDS with more detailed overview of work streams (RAS, public good, PIAG, private sector access framework). Scenario-based discussion in smaller groups to explore public good, public involvement and private sector access. Q&A with specialists in plenary.	RAS, public good and PIAG. Private sector access.
Online workshop Finalising conclusions and reflecting on concepts	Tuesday 20 May 2025, 6-9pm	All participants	Ratify conclusions on overarching questions and reflect on concepts	Playback of key findings and emerging conclusions. Response from RDS to emerging conclusions. Small group discussion to review and ratify conclusions.	Emerging conclusions. RDS response to emerging conclusions.

B. Recruitment

Overall, 36 participants were recruited to take part by telephone, using a screening questionnaire. The questionnaire captured demographic information about the participants, designed to help ensure the group's profile was broadly reflective of the Scottish population but also that those taking part in the in-person session could travel to Edinburgh. Quotas were set on age, gender, ethnicity, region and education. The aim was to achieve a sample of 36 participants (24 for the in-person route and 12 for the online only route), accounting for potential cancellations or drop-outs. In the end, 32 participants attended all the workshops (22 from the in-person route and 10 from the online route).

To support and enable participation in all workshops, and in line with industry standards, participants were each paid for their participation (those who took part online received £240 and those who took part online and in-person received £370 to account for the additional deliberation time as well as travel time and expense). At the recruitment stage, participants were provided with an information sheet and privacy notice (see below) containing details of the research and how their personal data would be used. Where necessary, participants were supported with training on how to use the technology and access the meeting platform for the online session. Workshops were also arranged to take place outside of regular office hours to increase participation.

The table below summarises the profile of participants taking part in the workshops against the quota targets set.

Recruitment quotas

Demographic	Variable	% in population[6]	Target (36)	Total recruited (36)	Total participating to end (32)
Gender	Women	51%	18	19	18
	Men	49%	18	17	14
Age	16-24	11%	4	4	3
	25-34	18%	6	6	5
	35-54	32%	12	12	11
	55+	38%	14	14	13
Ethnicity	Ethnic minority group	7%	4	5	4
	White	93%	32	31	28
Location	Central	12%	4	4	4
	Glasgow	13%	5	4	3
	Highlands & Islands	8%	5	4	3
	Lothians	15%	3	5	5
	Mid Scotland & Fife	12%	4	6	5
	North East Scotland	14%	5	4	3
	South	13%	5	4	4
	West	13%	5	5	5
Education	Level 4 (Degree/Professional qualification)	32%	12	12	10
	Level 3 (HNC/HND or equivalent)	13%	5	5	5
	Level 2 (Higher, A level or equivalent)	11%	4	9	9

---

[6] Population estimates are based on the Scottish 2022 Census where available, or mid-year estimates from National Records of Scotland (2021).

	Vocational qualification	8%	2	1	1
	Level 1 (O Grade, Standard Grade or equivalent)	19%	7	9	7
	No qualifications	17%	6	0[7]	0
		TOTAL	**36**	**36**	**32**

---

[7] Due to challenges meeting this target, focus was given to recruiting a higher number of participants with level 2 qualifications or below.

Participant recruitment questionnaire

24-092991-01 Researcher Access Service and Private Sector Access Framework:
Deliberative Workshops
Recruitment questionnaire v4

Dates/times:

- **Session 1 – Online webinar, Wednesday 23 April, 6-8.30pm**
- **Session 2A (in person) – Saturday 26 April, 10am-3pm. Venue is the Edinburgh Training and Conference Venue, 16 St Mary's Street, Edinburgh**
- **Session 2B (online) – Thursday 1 May, 6-9pm**
- **Session 3 – Online workshop, Tuesday 20 May, 6-9pm**

<INTRO>
Good morning/afternoon/evening. My name is ……. I'm calling on behalf of Ipsos Scotland, the independent research organisation to let you know about some new research that they are carrying out and to ask if you'd be interested in taking part.

Ipsos is carrying out this research on behalf of Research Data Scotland (RDS), who provide access to data about the Scottish public to approved researchers for scientific research. RDS is interested in understanding what the public think about how data is accessed for research.

[READ OUT FOR GROUP 1 (ONLINE/IN-PERSON) – CHECK LOCATION]

Participation involves being part of a group of 36 people who will meet over three workshop sessions in April and May this year. Two of these sessions will take place online with everyone, and one of them will take place in Edinburgh for those living nearby. During the sessions, you'll hear information about the topic from expert speakers, spend time discussing it with others and finding potential ways forward. By the end of the sessions, you and your fellow participants will have put together some key conclusions, and these will be presented to Research Data Scotland to inform their work. As a thank you for your time and to help cover any travel expenses, we would like to offer you £370.

[READ OUT FOR GROUP 2 (ONLINE ONLY) – CHECK LOCATION]

Participation involves being part of a group of 36 people who will meet over three workshop sessions in April and May this year. Two of the sessions will take place online with everyone, with the other either taking place in Edinburgh for those living nearby, or online for those living further away. During the sessions, you'll hear information about the topic from expert speakers, spend time discussing it with others and finding potential ways forward. By the end of the sessions, you and your fellow participants will have put together some key conclusions, and these will be presented to Research Data Scotland to inform their work. As a thank you for your time, we would like to offer you £240.

You don't need any prior knowledge to take part and there is no preparation you need to do. All that is required is a willingness to listen to the information presented and freely share your opinions.

Q1a. Is this something you would be interested in taking part in?

Yes	1	Continue
No	2	Thank and Close

Q1b. And are you available for all the dates? [if the person knows they can't or likely can't attend some sessions then this means they are not eligible; if they are worried that something might happen that would stop them attending then we can reassure them that they are eligible, and we understand that life happens, and we would cross that bridge if we come to it. The main thing is that, as it stands, they can make all those dates and will put them in their diaries!]

Yes	1	Continue
No	2	Thank and Close

GROUP 1 (ONLINE/IN-PERSON)

- **Session 1 – Online webinar, Wednesday 23 April, 6-8.30pm**
- **Session 2A (in person) – Saturday 26 April, 10am-3pm. Venue is the Edinburgh Training and Conference Venue, 16 St Mary's Street, Edinburgh**
- **Session 3 – Online workshop, Tuesday 20 May, 6-9pm**

GROUP 2 (ONLINE ONLY)

- **Session 1 – Online webinar, Wednesday 23 April, 6-8.30pm**
- **Session 2B (online) – Thursday 1 May, 6-9pm**
- **Session 3 – Online workshop, Tuesday 20 May, 6-9pm**

Q2. To help ensure the online events run smoothly, please can you let me know which of the following apply.

SINGLE CODE

a) I have broadband in my home which allows me to use video call apps and software to a good quality		
Yes	1	Continue
No	2	Discuss details and check with research team (we might be able to lend an internet dongle).
b) I have a laptop, PC or tablet which I can use for video calls		
Yes	1	Continue
No	2	Discuss details and check with research team (we

		might be able to lend a laptop).
c) I ….		
… have used Zoom before	1	Continue
… have not used Zoom before but I have used other video call apps and software	2	Continue (offer test Zoom call if that would be helpful)
… have not used it before but happy to give it a go with some direction	3	Continue and notify Ipsos team to arrange test call
… would not be comfortable taking part in the discussion on video software	4	Discuss details and offer one-to-one practice call with Ipsos (thank and close if participant still does not want to take part online)
d) I… [GROUP 1 ONLY]		
… am able to travel to a central venue in Edinburgh	1	Continue
… would not be comfortable taking part in-person	2	Discuss details and check with research team (may be able to flex as long as overall region spread is nat rep).

Q3. Could you tell me your postcode? [If needed: This is just to make sure we include people from across Scotland and can put you in the correct group for the middle workshop, which will either take place in-person if you live near Edinburgh, or online if you live further away]

Postcode

Q4. Do you or any members of your immediate family work in any of the following areas, either in a paid or unpaid capacity?
MULTICODE OKAY
READ OUT, MULTICODE

For Research Data Scotland (any role)		
For the Scottish Government (any role)		
For the University of Edinburgh (any role)		
Academic research, scientific research, or Research & Development		Thank and close (or collect more details and check with research team).
Advertising		

Data analytics or statistical research		
Data privacy law		
Journalism/the media		
No, none of these		Continue
Don't know		

Q5. Have you participated in a focus group discussion, or other research discussion for any market research company in the last 12 months?
SINGLE CODE

Yes	1	Thank and Close
No	2	Continue

Q6. Please could you tell me your age at your last birthday?

Record exact age

16-24	1	Recruit to quota – good mix of ages
24-34	2	Recruit to quota – good mix of ages
35-54	3	Recruit to quota – good mix of ages
55+	4	Recruit to quota – good mix of ages
Prefer not to say	5	Recruit to quota (discuss with research team for allocation to other quota group)

Q7. Which of the following best describes your gender?
SINGLE CODE

Man	1	Recruit to quota
Woman	2	Recruit to quota
Non-binary	3	Recruit to quota (then randomly reduce man/woman quotas to account for this)
My gender is not listed	4	Recruit to quota (then randomly reduce man/woman quotas to account for this)
Prefer not to say	5	Recruit to quota (discuss with research team for

		allocation to other quota group)

Q8. What is your ethnic group?
SINGLE CODE

White – Scottish White – British	1	Recruit to quota
White – Irish White – Gypsy/Traveller/Romany White – Polish White – Other (write in):	2	Recruit to quota (record ethnic group)
Mixed or multiple ethnic group (write in)	2	
Pakistani	2	
Indian	2	
Bangladeshi	2	
Chinese	2	
African	2	
Caribbean or Black	2	
Arab	2	
Other (Please write in)	2	
Prefer not to say	3	Recruit to quota (discuss with research team for allocation to other quota group)

Q9. What is the highest educational or professional qualification you have obtained?

SINGLE CODE

No formal qualification	1	Recruit to quota
‘O’ Grade, Standard Grade, GCSE, Intermediate 1, Intermediate 2	2	
Vocational qualification (SVQ1-2 or equivalent)	3	
Higher grade, A-levels, SVQ level 3 or equivalent	4	

HND, HNC, RSA Higher Diploma, SVQ Level 4-5 or equivalent	5	
First degree, higher degree or equivalent professional qualification	6	
Other	7	
Don't know	8	
Prefer not to say	9	

Q10a. And are you, yourself…
SINGLE CODE

Working 30 hours or more a week (Full time)	1	Record for analysis
Working 8 - 29 hours a week (Part-time)	2	
Not working (looking after home, unemployed, retired, student or other)	3	
Prefer not to say	4	

ASK IF WORKING
Q10b. What do you do for work?

Record job

Q11. And finally, which of the following statements comes closest to your view on sharing data about the public? (Rotate order between A and B with each call)

A - We should share all the public data we can because it benefits public services and me.

B - We should not share public data as the risks to people's privacy and security outweigh the benefits.

Agree much more with A than B	1	Record
Agree a little more with A than B	2	Record
Agree equally with both	3	Record

Agree a little more with B than A	4	Record
Agree much more with B than A	5	Record
Don’t agree with either	6	Record
Don’t know	7	Record
Refused	8	Record

Thanks very much, you are exactly the kind of person we are looking for. I now just need to inform you that the researchers will make an audio recording of the discussions. This is so that they can listen to them when they are doing their analysis. However, I would like to assure you that everything you say will be anonymous and will be treated confidentially. Is that ok?
[If no, please check with Ipsos team before confirming with participant]

Do you have any other accessibility requirements that we should be aware of to help make the event work for you? For example language, hearing, vision, or caring responsibilities? [If yes, please discuss with Ipsos team].

[READ OUT FOR GROUP 1 (ONLINE/IN-PERSON) – CHECK LOCATION]

The venue for the in-person workshop on Saturday 26 April, 10am-3pm is the Edinburgh Training and Conference Venue, 16 St Mary’s Street, Edinburgh.

Tea and coffee and lunch will be provided. Do you have any dietary requirements?

Record

As I mentioned, we would like to offer you £370 as a thank you for your time and to help cover any travel expenses. This will be paid in instalments after each session. You can choose to receive this either as cash (via a bank transfer) or as an e-voucher. Receiving a bank transfer for this amount may impact on any benefit payments or tax positions, so if you are concerned about this it is worth checking this before choosing a payment method – for example by speaking to your local Citizen’s Advice Bureau.

[READ OUT FOR GROUP 2 (ONLINE ONLY) – CHECK LOCATION]

As I mentioned, we would like to offer you £240 as a thank you for your time. This will be paid in instalments after each session. You can choose to receive this either as cash (via a bank transfer) or as an e-voucher. Receiving a bank transfer for this amount may impact on any benefit payments or tax positions, so if you are concerned about this it is

worth checking this before choosing a payment method – for example by speaking to your local Citizen's Advice Bureau.

EXPLAIN NEXT STEPS AND SEND PARTICIPANTS THE INFORMATION SHEET, PRIVACY NOTICE AND CONSENT FORM. ASK THEM TO READ THROUGH EVERYTHING AND RETURN SIGNED CONSENT FORM (EMAIL CONFIRMATION IS FINE) IF HAPPY TO PARTICIPATE.

RECORD PARTICIPANT INFORMATION (INCLUDING ANY ACCESSIBILITY REQUIREMENTS) IN SPREADSHEET AND HIGHLIGHT TO RESEARCH TEAM AS SOON AS POSSIBLE.

Participant information sheet (in-person)

Workshops on accessing data about the Scottish public for research

Information leaflet

Thank you for agreeing to take part in workshops on how data about the Scottish public is accessed for research.

Research Data Scotland want to understand what people living in Scotland think about how our data is accessed for research.

Ipsos is carrying out this research on behalf of Research Data Scotland, who provide access to data about the Scottish public to approved researchers for scientific research.

The purpose of the workshops is:

- to learn about and discuss how data is accessed for research
- to develop some conclusions to present to Research Data Scotland to inform their work.

You are one of 36 people from across Scotland who will meet over the next couple of months to learn about and discuss these issues.

Before taking part, it is important that you understand why this research is taking place and what it will involve. The main thing to remember is that you do not need any prior knowledge to take part!

Please read the information in this leaflet carefully and get in touch if you have any questions (see contact details at the end of the leaflet).

What is a deliberative workshop?

The type of workshop you will be part of is called a “deliberative workshop”. A deliberative workshop brings together a group of people to learn about a topic before discussing their views and drawing conclusions. The people who attend will listen to presentations from experts, learn about key issues, discuss them with one another, and then discuss and draw conclusions together.

When are the workshops?

We will meet for two online sessions and one in-person session across April and May. You need to be able to attend all three sessions, which will be:

- **Session 1 (online) – Wednesday 23rd April, 6-8.30pm**
- **Session 2 (in-person) – Saturday 26th April, 10am-3pm at the Edinburgh Training and Conference Venue, 16 St Mary’s Street, Edinburgh.**
- **Session 3 (online) – Tuesday 20th May, 6-9pm**

You will receive £370 as a thank you for taking part in the sessions. This will be paid in instalments.

You can choose to receive this either as cash (via BACS transfer) or in the form of an e-voucher. Receiving a bank transfer for this amount may impact on any benefit payments or tax positions, so it is worth checking this before choosing a payment method.

In-person workshop

The in-person workshop will be held at the Edinburgh Training and Conference Venue, 16 St Mary’s Street, Edinburgh, EH1 1SU. Tea and coffee and lunch will be provided. Please let us know if you have any dietary requirements or questions using the contact details on the final page of this information sheet.

REMEMBER! You don’t need any prior knowledge to take part; all we require from you is a willingness to listen to the information presented and share your opinions.

Who is carrying out this research?

Ipsos Scotland is an independent research agency carrying out this work on behalf of Research Data Scotland, who provide secure access to data about the Scottish public for research.

You can find out more about Ipsos Scotland here: https://www.ipsos.com/en-uk/scotland

You can find out more about Research Data Scotland here: https://www.researchdata.scot/

What information will be used?

The findings from the workshops will be anonymised, which means it will not be possible to identify anyone who took part.

There may be observers attending the workshop sessions. These will be people representing different organisations (for example, people from Research Data Scotland). We will inform you of these organisations and the reasons for their attendance in each session. This is a normal part of workshops like these and allows those with an interest to see how the process works. Observers will not contribute to or shape the discussion in any way but will simply listen in on some conversations and plenary sessions. There are strict controls on the number of observers for each session as well as guidelines that observers must agree to before attending.

When online presentations are given, these will be recorded so that they can be made available to you afterwards. Please note that you may be visible in these recordings unless you turn your webcam off. The chair will remind you about this.

If you have any queries about these aspects of the research, please get in touch (contact details on the next stage).

How will you record what we say?

We would like to audio-record our discussions (but will only do so with your agreement). Recording the discussion allows us to accurately capture what has been said.

How will the information be used?

Ipsos will write a report of the findings which will be used by Research Data Scotland to inform how they develop processes for providing secure access to public data for public sector, academic and private sector organisations.

The report is likely to use quotes, but we will not name or identify anyone when we use quotes. If you do not want us to quote anything you say, please just let us know.

Do I have to take part?

No. It is entirely your choice whether or not you take part. You can also choose not to answer individual questions and/or leave the discussion completely.

What are my rights under GDPR?

The General Data Protection Regulation (GDPR) is a set of rules and regulations designed to protect the privacy and personal data of individuals. We have a legal duty to inform you about why we are collecting information from you for the research.

We provide this information in our Privacy Notice which should be read alongside this information sheet. It outlines what information we will collect and why, and your rights.

What happens now?

If you are happy to take part, please complete the consent form slip and return it by email to the address below.

We really hope you enjoy being part of these important workshops!

Contacting us

For more information about the research, or if you have questions at any point, please use the contact details below.

✉ email us at:

☎ call or message us on:

Participant information sheet (online)

Workshops on accessing data about the Scottish public for research

Information leaflet

Thank you for agreeing to take part in workshops on how data about the Scottish public is accessed for research.

Research Data Scotland want to understand what people living in Scotland think about how our data is accessed for research.

Ipsos is carrying out this research on behalf of Research Data Scotland, who provide access to data about the Scottish public to approved researchers for scientific research.

The purpose of the workshops is:

- to learn about and discuss how data is accessed for research
- to develop some conclusions to present to Research Data Scotland to inform their work.

You are one of 36 people from across Scotland who will meet over the next couple of months to learn about and discuss these issues.

Before taking part, it is important that you understand why this research is taking place and what it will involve. The main thing to remember is that you do not need any prior knowledge to take part!

Please read the information in this leaflet carefully and get in touch if you have any questions (see contact details at the end of the leaflet).

What is a deliberative workshop?

The type of workshop you will be part of is called a “deliberative workshop”. A deliberative workshop brings together a group of people to learn about a topic before discussing their views and drawing conclusions. The people who attend will listen to presentations from experts, learn about key issues, discuss them with one another, and then discuss and draw conclusions together.

When are the workshops?

We will meet for three online sessions across April and May. You need to be able to attend all three sessions, which will be:

- **Session 1 – Wednesday 23rd April, 6-8.30pm**
- **Session 2 – Thursday 1st May, 6-9pm**
- **Session 3 – Tuesday 20th May, 6-9pm**

You will receive £240 as a thank you for taking part in the online sessions. This will be paid in instalments.

You can choose to receive this either as cash (via BACS transfer) or in the form of an e-voucher. Receiving a bank transfer for this amount may impact on any benefit payments or tax positions, so it is worth checking this before choosing a payment method.

REMEMBER! You don’t need any prior knowledge to take part; all we require from you is a willingness to listen to the information presented and share your opinions.

Who is carrying out this research?

Ipsos Scotland is an independent research agency carrying out this work on behalf of Research Data Scotland, who provide secure access to data about the Scottish public for research.

You can find out more about Ipsos Scotland here: https://www.ipsos.com/en-uk/scotland

You can find out more about Research Data Scotland here: https://www.researchdata.scot/

What information will be used?

The findings from the workshops will be anonymised, which means it will not be possible to identify anyone who took part.

There may be observers attending the online sessions. These will be people representing different organisations (for example, people from Research Data Scotland). We will inform you of these organisations and the reasons for their attendance in each session. This is a normal part of workshops like these and allows those with an interest to see how the process works. Observers will not contribute to or shape the discussion in any way but will simply listen in on some conversations and plenary sessions. There are strict controls on the number of observers for each session as well as guidelines that observers must agree to before attending.

When presentations are given, these will be recorded so that they can be made available to you afterwards. Please note that you may be visible in these recordings unless you turn your webcam off. The chair will remind you about this.

If you have any queries about these aspects of the research, please get in touch (contact details on the next stage).

How will you record what we say?

We would like to audio-record our discussions (but will only do so with your agreement). Recording the discussion allows us to accurately capture what has been said.

How will the information be used?

Ipsos will write a report of the findings which will be used by Research Data Scotland to inform how they develop processes for providing secure access to public data for public sector, academic and private sector organisations.

The report is likely to use quotes, but we will not name or identify anyone when we use quotes. If you do not want us to quote anything you say, please just let us know.

Do I have to take part?

No. It is entirely your choice whether or not you take part. You can also choose not to answer individual questions and/or leave the discussion completely.

What are my rights under GDPR?

The General Data Protection Regulation (GDPR) is a set of rules and regulations designed to protect the privacy and personal data of individuals. We have a legal duty to inform you about why we are collecting information from you for the research.

We provide this information in our Privacy Notice which should be read alongside this information sheet. It outlines what information we will collect and why, and your rights.

What happens now?

If you are happy to take part, please complete the consent form slip and return it by email to the address below.

We really hope you enjoy being part of these important workshops!

Contacting us

For more information about the research, or if you have questions at any point, please use the contact details below.

✉ email us at:

☎ call or message us on:

Participant privacy notice

This Ipsos UK Survey and your personal data

- 24-092991-01 RDS deliberative

This privacy notice explains who we are, the personal data we collect, how we use it, who we share it with, and what your legal rights are.

About Ipsos UK

- Ipsos MORI UK Limited and Ipsos (market research) Limited are a specialist research agency, commonly known as Ipsos and referred to in this privacy notice as "Ipsos UK". Ipsos UK is part of the Ipsos worldwide group of companies, and a member of the Market Research Society. As such we abide by the Market Research Society Code of Conduct and associated regulations and guidelines.

What is Ipsos UK's legal basis for processing your personal data?

- Ipsos UK requires a legal basis to process your personal data. Ipsos UK's legal basis for processing is your consent to take part in this research study. If you wish to withdraw your consent at any time, please see the section below covering 'Your Rights'.

How will Ipsos UK use any personal data including survey responses you provide?

- Firstly, responding to this study is entirely voluntary and any answers are given with your consent.
- Ipsos UK will keep your personal data and responses in strict confidence in accordance with this privacy notice and you will only be identifiable in any published results if you consent.
- We will only share recorded material(s) from the group with the Project Team; the Project Team is the Ipsos UK Project Team and supplier organisations working on this project.
- The recorded materials will be used for research and analysis purposes only.

Who we share your data with

- Ipsos UK will be using certain supplier organisations to assist us in running the research study and we will need to disclose your personal data to these supplier organisations for that purpose. These supplier organisations include:
- FieldMouse (recruitment agency)
- Verbit Go (for transcription of recorded in-person discussions)
- Large Language Models (Artificial Intelligence e.g. Open AI) using an Ipsos dedicated environment (for transcription of online interviews).

How will Ipsos UK ensure your personal information is secure?

- Ipsos UK takes its information security responsibilities seriously and applies various precautions to ensure your information is protected from loss, theft or misuse. Security precautions include appropriate physical security of offices and controlled and limited access to computer systems.
- Ipsos UK has regular internal and external audits of its information security controls and working practices and is accredited to the International Standard for Information Security, ISO 27001.

How long will Ipsos UK retain your personal data and identifiable responses?

- Ipsos UK will only retain your data in a way that can identify you for as long as is necessary to support the research project and findings. In practice, this means that once we have satisfactorily reported the anonymous research findings, we will securely remove your personal, identifying data from our systems.
- Your personal details used to invite you to take part in this market research study, will be securely deleted from our systems once the study and any quality control checks are complete; this is usually carried out within three months of project close.
- If you are identifiable in recorded focus group footage, it may be necessary to retain this material securely for a period of up to one year, after which it will be securely removed from our system.
- For this project we will securely remove your personal data from our systems by 07/2026 and from online hosting platforms by 01/2026.

Where will your personal data be held & processed?

- All of your personal data used and collected for this survey will be stored and processed in the United Kingdom and the European Union.

Your rights.

- You have the right to access your personal data within the limited period that Ipsos UK holds it.
- Providing responses to this survey is entirely voluntary and is done with your consent. You have the right to withdraw your consent at any time whilst we hold your personal data at an identifiable level.
- You also have the right to request from us the deletion or erasure of the personal information we hold about you.
- You also have the right to rectify any incorrect or out-of-date personal data about you which we may hold.
- If you want to exercise your rights, please contact us at the below Ipsos UK address.
- If you have any complaints, we will appreciate if you give us an opportunity to resolve any issue first, by contacting us as set out below. You are, however, always entitled to contact the UK's Information Commissioner at Make a complaint | ICO.

How can you contact Ipsos UK about this survey and/or your personal data?

Email: <Ipsos research team email address> with "24-092991-01 RDS deliberative" in the email subject line

Post: 24-092991-01 RDS deliberative
Data Protection Officer, Compliance Department
Ipsos (market research) Limited and Ipsos MORI UK Limited
3 Thomas More Square
London E1W 1YW
United Kingdom

C. Discussion guides

RDS deliberative

Workshop 1: Learning webinar

Session 1: Wednesday 23rd April 2025, 6-8.30pm
Group of 36 participants, with 4 pre-assigned break-out groups (of 9 participants each)
Overarching objectives:

- Participants are introduced to each other and become familiar with the process and topic area.
- Opportunity to share initial thoughts and perceptions on: public sector data and how it is accessed for research; public good; the role of RDS, and what previous public engagement has found.

Time allocated	Discussion structure	Objectives	Questions and materials
17.30 – 17.50	Set-up: Facilitators check in 20 mins	Ensure technology is set-up correctly before participants enter the room	**Chair, facilitators and tech team only** • Test link, mics and cameras. • Test who has the host/co-host function and ensure it is allocated to the right team member(s) for assigning break out rooms. • Make all facilitators co-hosts. • Change screen name to NAME – Org – Chair/Facilitator. • Check everyone is on the WhatsApp group for facilitation team to be able to ask questions • Facilitators, speakers and observers allowed in early and asked to change screen name • Meanwhile tech support is assigning participants who are in the waiting room, facilitators, experts and observers to break-out rooms.
17.50 – 18.00	Participant check-in 10 mins	Enable participants to get settled and resolve any tech issues	**Participants log into the online session** • Participants encouraged to join the Zoom session early to check in and check their video/mic. • Participants encouraged to get a pen and paper and get settled somewhere quiet. • Register as people join and change screen names as necessary to first name and first initial of surname (e.g. Jane S).

Time allocated	Discussion structure	Objectives	Questions and materials
18.00 – 18.15	Welcome, introductions and context setting 15 mins	To welcome and orientate participants, enabling them to settle in and providing a sense of safety	Participants allocated to break-out groups, but not put in them. **Ipsos Chair to welcome everyone to the session:** • Chair welcomes participants to this first workshop on how data about the Scottish public is accessed for research. • Chair introduces poll and asks participants to answer the following question: How much (if anything) would you say you know about the use of data about people in Scotland for research? 1. A great deal 2. A fair amount 3. Just a little 4. Nothing at all 5. Don't know • Explains overall aim of the research, what their participation will involve, and the role that the public will play. Highlights what the end goal is i.e. to help RDS to develop processes to provide secure access to data that is not identifiable for public sector, academic and private sector organisations. **1. How can we tell when the use of de-identified data about people in Scotland for research is in the 'public good', and when it is not?** Chair to explain that de-identified data means details like names, birth dates and addresses are removed so researchers can never identify an individual. **2. How should the public be involved in decisions about sharing de-identified data about people in Scotland?** **3. Which methods for private sector organisations accessing data about people in Scotland are acceptable?** • Provides summary of overall process (i.e. number of future workshops) and today's agenda (including time of breaks and finishing time). Explain that today's session will mostly be about

Time allocated	Discussion structure	Objectives	Questions and materials
			listening and learning and encourage participants to jot down their thoughts and questions, explaining that there will be opportunities for Q&A. Encourage participants to grab a pen and paper, or a device that they can jot notes down on if they wish. • Explains who is here – our group of participants representing people from across Scotland, Ipsos facilitators, presenters, and any observers. • Housekeeping, ground rules – mention that plenary sessions will be recorded so keep camera off if don't want to be visible during that. Reminder to only have first name and first letter of surname showing. • Provide assurances around the scope of the research and confidentiality
18.15 Move to pre-allocated break-out rooms			
18.15 – 18.25	Break-out group introductions and warm-up discussion 10 mins	To allow participants to meet others and warm up (if in group) / explore top-of-mind perceptions of data sharing	**Break-out group introductions (10 mins)** Facilitator welcomes participants, thanks them for their participation and introduces themselves. Confirms permission for recording. • Please briefly tell everyone your first name, what part of Scotland you live in, and share why you decided to take part in this research. Warm up exercise: • We'll start by thinking about data. Note down 2 or 3 words that spring to mind when you think about data? ○ Let's go around and see what everyone put, and why • What did you make of the question we asked at the start? • Did the results surprise you? Why/why not? • [If time] Do you have any initial thoughts on the use of data about people in Scotland for research? Note down responses to refer back to later ○ What are the benefits? ○ What are the risks? ○ IF ANYONE RESPONSE 'IT DEPENDS', ask them to elaborate (what does it depend on, why is that the case?)

Time allocated	Discussion structure	Objectives	Questions and materials
			Facilitator thanks participants, and acknowledges we will now move to plenary to see the first presentation.
18.25 move to plenary			
18.25 – 18:35	Learning Stage Presentation 1 10 mins	Presentation to inform participant learning	**Plenary presentation 1** (live) (10 mins) – **Introduction to key concepts** • What we mean by 'data about people living in Scotland' ○ Deidentified person level administrative data • How data used are used, by whom, in research • How data are protected, e.g.: ○ Five Safes Framework, particularly: ▪ Safe Data –emphasise deidentified data (details like names, birth dates and addresses are removed so researchers can never identify an individual.) ▪ Safe People – ensuring we keep this open on private sector and public sector ○ TREs, Safe Havens ▪ Keep focus on national safe haven • What are the benefits of using data in research? • What are the risks and challenges?
18.35 move to break-out rooms			
18.35 – 18.50	Reflections 15 mins	To enable participants to reflect and gather initial views, and form any questions	**Reflections on presentation 1 (10 mins)** If needed: remind/reassure participants that there is a lot of information being shared in this session, and that recordings will be shared afterwards so they can reflect on things in their own time. Facilitator can bring slides up if helpful for participants to see anything again. • What are your reflections on what you've just heard? • Was there anything that surprised you or stood out to you? • [Probes if needed] And what did you think about: ○ The different types of data about people in Scotland? ○ How data about people in Scotland are used? ○ How data are 'de-identified'? [If needed – 'this is when identifiable information, such as names and addresses, are removed from the data'] ○ How data are protected (e.g. the Five Safes, Safe Havens, TREs)? ○ The benefits of using data about people in Scotland in research?

Time allocated	Discussion structure	Objectives	Questions and materials
			○ The risks and challenges of using such data in research? • [If applicable] Was there anything that you learned that has changed your views from earlier [on access to data about people in Scotland for research]? • Was anything unclear? • Was there anything you want to hear more about? **Question forming (5 mins)** • What questions do you want to ask [speaker]? • What is our priority question? And what about a back-up in case our question is asked by another group? • Does anyone want to ask our question on behalf of the group?
18.50 move to plenary			
18:50 – 19:00	Learning Stage Presentation 2: 10 mins	Presentation to inform participant learning	**Plenary presentation 2** (live) (10 mins) - **Public good** • Overview or data ethics, what it is and why it matters • What does 'public good' mean • The challenge of defining 'public good' (reiterating how this group can help with that) • What we know about public views on public good (reiterating how this research will build on that)
19.00 move to break-out rooms			
19:00 – 19:15	Reflections 15 mins	To enable participants to reflect and gather initial views, and form any questions	**Reflections on presentation 2 (10 mins)** If needed: remind/reassure participants that there is a lot of information being shared in this session, and that recordings will be shared afterwards so they can reflect on things in their own time. Facilitator can bring slides up if helpful for participants to see anything again. • What are your reflections on what you've just heard? • Was there anything that surprised you or stood out to you? • [Probes if needed] And what did you think about: ○ Data ethics ○ The idea of 'public good' ○ The challenge of defining 'public good' ○ What we know about public views on public good

Time allocated	Discussion structure	Objectives	Questions and materials
			• [If applicable] Was there anything that you learned that has changed your views from earlier [on access to data about people in Scotland for research]? • Was anything unclear? • Was there anything you want to hear more about? **Question forming (5 mins)** • What questions do you want to ask [speaker]? • Does anyone want to ask our question on behalf of the group?
19.15 - 19.25 Break			
Move to plenary 19.25			
19.25-19.40	**Q&A**	To allow participants to pose questions to speakers	Plenary (15 mins): chair to facilitate Q&A session, with facilitators asking the questions or calling on participants to ask their question.
Stay in plenary			
19.40 – 19.50	Learning Stage Presentation 3: 10 mins	To enable participants to reflect and gather initial views, and form any questions	Plenary presentation 3 (live) (10 mins) - Overview of RDS, and brief intro to Researcher Access Service, and Private Sector Access Framework (reiterating role of participants going forward) • What RDS is • Why RDS exists • What RDS has been doing to date and plans for the future • Why RDS wants to hear from the public • Brief introduction to Researcher Access Service and Public Involvement Advisory Group • Brief introduction to Private Sector Access Framework
19.50 move to break-out rooms			
19.50 – 20.05	Reflections 15 mins	To enable participants to reflect and gather initial	**Reflections on presentation 3 (10 mins)** If needed: remind/reassure participants that there is a lot of information being shared in this session, and that recordings will be shared afterwards so they can reflect on things in their own time. Facilitator can bring slides up if helpful for participants to see anything again.

Time allocated	Discussion structure	Objectives	Questions and materials

		views, and form any questions	• What are your reflections on what you've just heard?
			• Is there anything that surprised you?
			• [Probes if needed] And what did you think about:
			  ○ What RDS do?
			  ○ The Researcher Access Service
			  ○ The role of the general public
			  ○ Private Sector Access Framework
			• [If applicable] Was there anything that you learned that has changed your views from earlier [on access to data about people in Scotland for research]?
			• Was anything unclear?
			• Was there anything you want to hear more about?
			
			Question forming (5 mins)
			• What questions do you want to ask [speaker]?
			• Does anyone want to ask our question on behalf of the group?

Move to plenary 20.05			
20.05 – 20.15	**Q&A**	To allow participants to pose questions to speakers	Plenary (10 mins): chair to facilitate Q&A session, with facilitators asking the questions or calling on participants to ask their question.
20.15 move to breakouts			
20.15 – 20.25	Initial conclusions 10 mins	Consolidate learning, form initial conclusions	Initial conclusions (10 mins) Facilitator have Miro board open ready to share screen. We learned about public good this evening. One of the questions you are tasked with answering is "How can we tell when the use of de-identified data about people in Scotland for research is in the 'public good', and when it is not?" Post-it exercise • We're now going to summarise our thoughts on that, based on everything we've heard so far, by noting them down on a virtual post-it note (facilitator share screen). On the screen here you can see a bunch of coloured squares - our group's are all coloured (COLOUR).

Time allocated	Discussion structure	Objectives	Questions and materials
			• I want each of you to think about how you would complete this sentence - "To me, using data for research is in the public good if…" • So if you have a think, I will write it on the square – one per person. • We can also see what the other groups are writing, and we will try to group them together if they are similar IF THERE IS TIME • Finally, before we report back, what has stood out for you most this evening?
20.25 move to plenary			
20.25 – 20:30	Summary and close 5 mins	To orientate participants and encourage continued engagement	Opportunity for sharing before chair closes session • Brief feedback from each group on the key reflections/issues coming out so far Chair to close the day: • Brief overview of what has been covered • Brief overview of what to expect in later workshops • Encourage participants to speak to family and friends and see if they have similar/different views on the key issues • Chair highlights short survey participants asked to complete after this workshop • Reminder of incentives process **Thank participants and close** Facilitators and tech support will only leave once all participants have left, and all facilitators can encourage participants to come off mute, wave and say goodbye by doing so. Facilitators to make a note of any questions raised in group discussions that weren't addressed in the Q&A sessions ready to share with research team.
			Short survey questions for after the workshop Q1. How much (if anything) would you say you know about the use of data about people living in Scotland for research? SINGLE CODE ONLY

Time allocated	Discussion structure	Objectives	Questions and materials
			1. A great deal 2. A fair amount 3. Just a little 4. Nothing at all 5. Not sure 6. Prefer not to say Q2a. How well would you say you understand how data is protected? SINGLE CODE ONLY 7. Very well 8. Fairly well 9. Not very well 10. Not well at all 11. Not sure 12. Prefer not to say Q2b. How well would you say understand the concept of ‘public good’ when using data about people living in Scotland for research ? SINGLE CODE ONLY 1. Very well 2. Fairly well 3. Not very well 4. Not well at all 5. Not sure 6. Prefer not to say Q2c. How well would you say understand the purpose of Research Data Scotland (RDS)? SINGLE CODE ONLY 1. Very well 2. Fairly well 3. Not very well 4. Not well at all 5. Not sure

Time allocated	Discussion structure	Objectives	Questions and materials
			6. Prefer not to say Q5. Do you have any other feedback you'd like to share? OPEN QUESTION

RDS deliberative

Workshop 2a: in-person deliberation

Session 2a: Saturday 26th April 2025, 10am-3pm

Group of 24 participants, with 3 pre-assigned groups (of 8 participants each)

Overarching objectives:

- Participants explore the topics in more detail, using scenarios to consider: the definition of public benefit for non-health related research; what public involvement should look like in the Researcher Access Service; what criteria for allowing private sector access to public sector data would be acceptable.
- Participants reflect, discuss and deliberate to reach initial conclusions to the three questions.

Time allocated	Discussion structure	Objectives	Questions and materials
09.30 – 10.00	**Set-up and registration** **30 mins**	Set-up	• Facilitators to set up room (3 tables, screen for connecting to laptop and showing slides, refreshments, handouts, stationery at tables) • Check on fire-exits, location of toilets, and any access requirements • On arrival, participants will be given a name badge and sent to their pre-allocated table
10-10.10	**Introduction and scene setting** **10 mins**	Making participants feel welcome and informed	Ipsos Chair to give a warm welcome to everyone and thank them for returning to the second workshop (10 mins): • Chair to remind participants of the overall aims and purpose of the workshops (including reminder of who Ipsos and RDS are) and the role that participants are playing. • Recap who is in the room, including introducing any new speakers/observers who were not present at Workshop 1. • Reminder what they discussed last time (signposting to the poster with post-its on what public good means covered in WS1) and the questions they will answer by the end of the workshops: ○ **How can we tell when the use of de-identified data about people in Scotland for research is in the ‘public good’, and when it is not?** ○ **How should the public be involved in decisions about sharing de-identified data about people in Scotland?** ○ **Which methods for private sector organisations accessing data about people in Scotland are acceptable?**

Time allocated	Discussion structure	Objectives	Questions and materials
			• Run though the agenda and what to expect from workshop 2. Emphasising that this time there will be less listening and learning, and more discussion from the participants themselves. They will be given a quick reminder of some of the key points from last time, and then will spend time looking at different scenarios and examples that should help to bring the topic to life. • Explain confidentiality and anonymity, housekeeping (including fire exits and toilets), ground rules. • Explanation/ reminder of some of the key concepts they learned about in session 1, to support their discussions today (recapping on key concepts like data about people living in Scotland, 5 Safes Framework, public good, RDS, Researcher Access Service, Private Sector Access Framework). A glossary of key terms will be available on the tables and facilitators will encourage participants to refer to these when needed. • Explain that some of the scenarios we will be discussing today include some sensitive information or may cover situations that affect you or people you know. Show the wellbeing signposting slide and explain that these are some support links in case any of the discussion today raises things that are upsetting. Also reassure participants that it is fine to step away from the discussion if you need a breather, just let your facilitator know.
10.10 Breakouts			
10.10-10.15	Introductions Breakout 1 5 mins	Warm up and checking in	**Introductions (5 mins)** Facilitator introduces themselves and thanks for participating. Collects permission to audio record the discussion and then asks participants to introduce themselves. • What has stuck in your mind most since the first session? • Have any new questions or issues come up for you since the last session? • How do you feel about the plan for today?
10.15 Plenary			

Time allocated	Discussion structure	Objectives	Questions and materials
10.15-10.35	Presentation to inform participant learning 20 mins	Further information/context setting	**Plenary presentation 1 (live) (20 mins) - The Researcher Access Service and public good** (including what past previous engagement has found and how it has informed the RAS to date, reiterating how this group will inform the next stage of development) • Brief recap on who RDS is • Introduce Researcher Access Service – what it is, how it came about, why it's needed, who is involved in the RAS panel • How public engagement has informed the RAS to date • Introduce Public Impact Advisory Group pilot ○ Background (i.e. why/how the pilot came about) ○ What it will look like ○ When/how it could be involved in decision-making ○ Recapping how this group will inform the next stage of development ○ Broaden access to different types of data, not just health ○ Inform the decision making of RAS Panel ○ Inform set up of PIAG • Walk through of a health project example and the public good criteria that is currently being used for access to health data for research, how it was assessed, and how it is tracked/reported on)
Breakouts			
10.35 – 10.50	Reflections 15 mins	To enable participants to reflect and gather initial views, and form any questions	**Reflections on presentation 1 (10 mins)** • What are your reflections on what you've just heard? • Was there anything that surprised you or stood out to you? • [Probes if needed] And what did you think about: ○ The recap on RDS and its role/purpose? ○ The Researcher Access Service and process for accessing data ○ The Public Impact Advisory Group and its role • Was anything unclear? • **Question forming (5 mins)** • What questions/comments do you want to ask/make to [speaker]? • Does anyone want to ask our question on behalf of the group?

Time allocated	Discussion structure	Objectives	Questions and materials
10.50 – 11.00	**Q&A** **10 mins**		Plenary (10 mins): chair to facilitate Q&A session, with facilitators asking the questions or calling on participants to ask their question. Chair to remind participants that experts can be called over to tables at any points throughout the day to answer questions.
BREAK 11.00 – 11.10			
Breakouts			
11.10 – 12.00	**Group discussions and deliberation on RAS and public involvement**	Scenario-based discussion to start addressing Q1 (How can we tell when the use of data about people in Scotland for research is in the 'public good', and when it is not?) Discussion to start addressing Q2 (How should the public be involved in decisions about sharing data about people in Scotland?)	**Scenarios (40 mins)** We're now going to look at some scenarios… Facilitator share handout and read through research summary first. Check participants are clear about it / clarify any questions as needed. Scenario order (Aim for 2 scenarios): Facilitator 1: A,B,C Facilitator 2: B,C,A Facilitator 3: C,A,B Read through scenarios in order and, for each, ask: • Who is requesting access? • What type of data do they need? • Why do they need it? • What impact(s) could it have? • What are the potential benefits? • Who could benefit? IF NEEDED, PROBE ON: ○ Which individuals? How? ○ Which groups? How? ○ Or all of Scotland? How? ○ How big would the benefit(s) be? • What risks are there? IF NEEDED, PROBE ON: ○ Stigmatising certain groups? ○ Unfair treatment of/decisions about certain groups? ○ Bias? ○ Representativeness and accuracy of decisions? ○ Financial? ○ How big would that risk/those risks be?

Time allocated	Discussion structure	Objectives	Questions and materials
			○ Who could be affected if these risks became reality? How would they be affected? • ASK ABOUT BENEFITS/RISKS SPONTANEOUSLY FIRST THEN SHOW POTENTIAL RISKS/BENEFITS ON OTHER SIDE OF SCENARIO AND GATHER THOUGHTS ON THOSE. • What other information do we need to determine whether or not this project is for a public good? Why is that important? • If RDS was to provide this type of data to organisations like this one, do you think the benefits outweigh the risks? Why/why not? ○ [If not] What would need to change to ensure the benefits outweigh the risks? Facilitator note down benefits and risks as mentioned by participants (use different pen colours to distinguish between scenarios). Now if we assume that this project had been approved and has gone ahead... • How would we check that the project had done what it said it would? ○ What, if any, tracking or monitoring should be in place? ○ Currently, all approved projects are published on the RDS website. RDS are considering asking organisations to report back to them after a certain amount of time or at the end of the project. What are your thoughts on that approach? ▪ What kind of things would you want to know about? ▪ How much time after should they report back? (if response is “it depends, probe on what it would depend on in this scenario) • Facilitator recap on the PIAG [Public Impact Advisory Group which RDS would like to set up to involve members of the public in discussions about applications before they go to the approvals panel], then ask: ○ What would the benefits be of involving the public in decisions about this project? ○ What would the drawbacks be? ○ On balance. do you think the Public Impact Advisory Group should be involved in decisions about this project? ▪ If no, why not? ▪ If yes, why? • Is there any other information that you think members of the Advisory Group would need to help them review a project application? Facilitator note down key themes on public involvement mentioned by participants (on new sheet), using new colour pen to distinguish between scenarios.

Time allocated	Discussion structure	Objectives	Questions and materials
			Throughout discussion, facilitator to play back emerging themes that address the question of determining whether a data use request is for public good or not, and public involvement, noting down on flipchart (one sheet for public good and a separate sheet for PIAG). **Post-it exercise (10 mins)** Now, thinking about all the scenarios we've looked at, I want you to think about how you would assess the public good of these data requests and what your conclusion would be. There is a poster over there with 4 spaces, or quadrants, and I want you to take a post-it, write the letter of the scenario and the name of the project on the post-it, and then go place it on the poster where you feel it should go. The bottom left of the poster means the project has a low risk and low benefit, the bottom right means there is low risk and high benefit, the top right means high risk and high benefit, and the top left means there is high risk and low public benefit. So you can place your post-it anywhere (e.g. if you're between two spaces you would place it along the line between them). Facilitator hands out post its and signposts participants to the poster with the benefit/risk quadrant where they can place their post its according to how they feel about the overall public good v risks. Facilitator encourages participants to look at the post-its from WS1 to think about what public good means.
LUNCH BREAK: 12.00 – 12.45 (45 MINS)			
Plenary			
12:45 – 13.05	**Feedback followed by presentation to inform participant learning** **20 mins**	Further information/context setting	**Brief feedback (5 mins)** Before introducing next presentation, Chair to briefly comment on the benefit/risk quadrant with participants' post-its added, and what it is showing, inviting facilitators to briefly summarise feedback in relation to one scenario their group looked at. **Plenary presentation 2 (live) (15 mins) - Private Sector Access Framework** • Background – why is an approach to private sector access needed • RDS position on private sector access • Give overview of opportunities and risks with private sector involvement • Introduce PSAF – what it is and work to date • Introduce models of private sector involvement

Time allocated	Discussion structure	Objectives	Questions and materials
			- Reiterate role of group (i.e. to consider these models and share views on what is acceptable/unacceptable in relation to private sector access to data)
Breakouts			
13.05 – 13.15	Reflections 10 mins	To enable participants to reflect and gather initial views, and form any questions	**Reflections on presentation 2 (5-8 mins)** • What are your reflections on what you've just heard? • Was there anything that surprised you or stood out to you? • [Probes if needed] And what did you think about: ○ The opportunities that private sector involvement could create? ○ The risks that private sector access could create? ○ The Private Sector Access Framework • How are you currently feeling about data sharing between public organisations and private organisations? • **Question forming (2-5 mins)** • What questions/comments do you want to ask/make to [speaker]? • Does anyone want to ask our question on behalf of the group?
Plenary			
13.15 – 13.25	Q&A	To allow participants to pose questions to speakers	Plenary (10 mins): chair to facilitate Q&A session, with facilitators asking the questions or calling on participants to ask their question. Chair to remind participants that experts can be called over to tables at any points throughout the day to answer questions.
Breakouts			
13.25– 14.15	**Group discussions and deliberation on private sector access public involvement**	Scenario-based discussion to start addressing Q3 (When is it acceptable and when it is unacceptable for private sector organisations to	**Scenarios (50 mins)** We're now going to look at some scenarios… Facilitator hand out pack with scenarios and model summaries. Read through cover sheet first and check understanding. Scenario order (aim for at least one scenario covering each model): Facilitator 1 – F,D,E Facilitator 2 – D,E,F Facilitator 3 – E,F,D

Time allocated	Discussion structure	Objectives	Questions and materials
		access data about people in Scotland?) Discussion to start addressing Q2 (How should the public be involved in decisions about sharing data about people in Scotland?)	Model 1: Research for, but not by, a private sector company Model 2: Research by a private sector company working with others in the public sector Model 3: Research by a private sector organisation not in partnership Model order: Facilitator 1 – 1,2,3 Facilitator 2 – 3,1,2 Facilitator 3 – 2,3,1 Facilitator to talk through scenario and ask for initial views: • Optional probes if participants need some time to familiarise (otherwise skip): o Who is requesting access? o What type of data do they need? o Why do they need it? • What impact(s) could this project have? • Who could benefit from it? IF NEEDED, PROBE ON: o Which individuals? How? o Which groups? How? o Or all of Scotland? How? o How big would that impact/those impacts be? • How big would that impact/those impacts be? • What risks are there? o Stigmatising certain groups? o Unfair treatment of/decisions about certain groups? o People’s privacy? o Bias? o Representativeness and accuracy of decisions? o Financial? o Legal? o How big would this risk/those risks be? • Who could be affected if these risks became reality? How would they be affected? Then take each model in turn, read through, and ask: • What are your thoughts on this model? o What do you like about it? o What do you dislike about it?

Time allocated	Discussion structure	Objectives	Questions and materials
			• What would this model mean for the data access request we are looking at? ○ Who would access the data? ○ [For models 1 and 2] What public sector organisations, if any, would be interested in partnering with the company? ○ How would it be used? • To what extent, if at all, would this model ensure the benefits we just discussed? Probe fully on how/reasons why not • To what extent, if at all, would this model address the risks we just discussed? Probe fully on how/reasons why not • Would it be acceptable or unacceptable to you if RDS was to allow [organisation] access to this data under this model? Why? ○ [If unacceptable] What would need to be in place to make it acceptable? • Now if we assume that this project had been approved and has gone ahead... How would RDS make sure that the project is doing what it said it would and share that information? • Does this bring up any new or different considerations for how we would consider the 'public good' of this project? Facilitator note down benefits, risks and thoughts on models as mentioned by participants (on new sheet). After discussing all models in relation to scenario, ask: • Thinking about this scenario, which model do you think would work best? Why? • And which model do you think would work least? Why? And now coming back to public involvement. The Public Impact Advisory Group is something that RDS would like to set up to involve members of the public in discussions about applications before they go to the approvals panel: ○ What would the benefits be of involving the public in decisions about this project? ○ What would the drawbacks be? ○ On balance, do you think the Public Impact Advisory Group should be involved in decisions about this project? ▪ If no, why not? ▪ Is this the same or different to the scenarios we looked at earlier involving public sector organisations? DEPENDING ON RESPONSE, PROBE: • What makes it the same? • What makes it different?

Time allocated	Discussion structure	Objectives	Questions and materials
			Facilitator note down key themes on public involvement mentioned by participants (using previous public involvement sheet, using new colour pen to distinguish between scenarios). REPEAT FOR NEXT SCENARIO IF TIME Throughout discussion, facilitator to play back emerging themes that address the question of when private sector access to data is acceptable or not, and public involvement noting down on flipchart. Use one sheet for private sector access, and add to existing sheet on public involvement (use different colour pen to denote themes in relation to private sector). Chair invites participants to stretch legs for 5 mins, grab a drink etc. Encourage participants to look at the risk/benefit quadrant and 'public good' post-it poster ahead of final exercise.
BREAK: 14.15 – 14.20 (5 MINS)			
Breakouts			
14.20 – 14.50	**Initial conclusions** **30 mins**	Summarising initial responses to key overarching questions	In this last section we're going to bring together some initial conclusions to those three key questions that we showed you at the beginning. Question order: Facilitator 1 – 1, 2, 3 Facilitator 2 – 3, 2, 1 Facilitator 3 – 2, 1, 3 1. **How can we tell when the use of de-identified data about people in Scotland for research is for the 'public good', and when it is not?** Facilitator refer back to flipchart/notes and ask if there's anything to add, change or expand on (asking participants to explain their reasons and asking others in the group whether they agree or disagree – reassuring that we don't have to reach a consensus, but that we want to make sure that everyone has had their say). Encourage participants to revisit the benefit/risk quadrant. Additional probes if needed: • Thinking back on the scenarios we looked at earlier today (briefly recap/summarise/show flipchart), what other issues did they raise that might be relevant to this question? • What if this data access request didn't come from [public sector organisation], but a private sector organisation instead? Would that change things or not? Why not/how so? ○ Do we need to add anything to our 'public good' criteria to make sure this is covered?

Time allocated	Discussion structure	Objectives	Questions and materials
			- [If time] [The scenario] was placed [summarise placement of scenario on chart] – why do you think that was the case? What does that mean for determining whether or not a project is for a 'public good'? 2. **How should the public be involved in decisions about sharing de-identified data about people in Scotland?** Facilitator refer back to flipchart/notes and ask if there's anything to add, change or expand on (asking participants to explain their reasons and asking others in the group whether they agree or disagree – reassuring that we don't have to reach a consensus, but that we want to make sure that everyone has had their say) - Should public involvement look the same or different for different types of research project? Why/why not? Facilitator refer back to RAS scenarios and types of projects (e.g. types of data, types of organisation) - Should public involvement look the same or different if the private sector was involved in a research project? Why/why not? Facilitator refer back to private sector scenarios and themes in relation to public involvement on private sector 3. **Which methods for private sector organisations accessing data about people in Scotland are acceptable?** Facilitator refer back to flipchart/notes and ask if there's anything to add, change or expand on (asking participants to explain their reasons and asking others in the group whether they agree or disagree – reassuring that we don't have to reach a consensus, but that we want to make sure that everyone has had their say) Additional probes if needed: - Thinking back on the scenarios/models we looked at earlier today (briefly recap/summarise), what other issues did they raise that might be relevant to this question?
Plenary			

Time allocated	Discussion structure	Objectives	Questions and materials
14.50 – 15.00	**Wrap-up** **10 mins**		Opportunity for sharing before chair closes session • Brief feedback from each group on initial answers to the questions. Chair to close the day: • Brief overview of what has been covered • Brief overview of next steps i.e. similar workshop with online group; RDS will reflect on the feedback so far and share some draft conclusions at the final online workshop • Chair highlights short survey participants asked to complete after this workshop • Reminder of incentives process **Thank participants and close** Facilitators and support will only leave once all participants have left. Facilitators to make a note of any unanswered questions to share with research team.

RDS deliberative

Workshop 2b: online deliberation

Session 2b: Thursday 1st May 2025, 6pm - 9pm

Group of 12 participants, with 3 pre-assigned groups (of 4 participants each)

Overarching objectives:

- Participants explore the topics in more detail, using scenarios to consider: the definition of public benefit for non-health related research; what public involvement should look like in the Researcher Access Service; what criteria for allowing private sector access to public sector data would be acceptable.
- Participants reflect, discuss and deliberate to reach initial conclusions to the three questions.

Time allocated	Discussion structure	Objectives	Questions and materials
17.30 – 17.50	**Set-up: Facilitators check in** **20 mins**	Ensure technology is set-up correctly before participants enter the room	**Chair, facilitators and tech team only** • Test link, mics and cameras. • Test who has the host/co-host function and ensure it is allocated to the right team member(s) for assigning break out rooms. • Make all facilitators co-hosts. • Change screen name to NAME – Org – Chair/Facilitator. • Check everyone is on the WhatsApp group for facilitation team to be able to ask questions • Facilitators, speakers and observers allowed in early and asked to change screen name • Meanwhile tech support is assigning participants who are in the waiting room, facilitators, experts and observers to break-out rooms.
17.50 - 18.00	**Participant check-in** **10 mins**	Enable participants to get settled and resolve any tech issues	**Participants log into the online session** • Participants encouraged to join the Zoom session early to check in and check their video/mic. • Participants encouraged to get a pen and paper and get settled somewhere quiet. • Register as people join and change screen names as necessary to first name and first initial of surname (e.g. Jane S).
18.00 - 18.10	**Welcome, and context setting** **10 mins**	To welcome and orientate participants, enabling them to settle in and	Participants allocated to break-out groups, but not put in them **Ipsos Chair to welcome everyone to the session:** • Chair gives everyone a warm welcome and thanks them for returning to the second workshop.

Time allocated	Discussion structure	Objectives	Questions and materials
		providing a sense of safety	• Chair to remind participants of the overall aims and purpose of the workshops (including reminder of who Ipsos and RDS are) and the role that participants are playing. • Recap who is here, including introducing any new speakers/observers who were not present at Workshop 1. • Reminder what they discussed last time (signposting to the post-its on what public good means covered in WS1 which were emailed to participants before session) and the questions they will answer by the end of the workshops: ○ **How can we tell when the use of de-identified data about people in Scotland for research is in the 'public good', and when it is not?** ○ **How should the public be involved in decisions about sharing de-identified data about people in Scotland?** ○ **Which methods for private sector organisations accessing data about people in Scotland are acceptable?** • Run though the agenda and what to expect from workshop 2. Emphasising that this time there will be less listening and learning, and more discussion from the participants themselves. They will be given a quick reminder of some of the key points from last time, and then will spend time looking at different scenarios and examples that should help to bring the topic to life. • Explain confidentiality and anonymity, housekeeping, ground rules. • Explanation/ reminder of some of the key concepts they learned about in session 1, to support their discussions today (recapping on key concepts like data about people living in Scotland, 5 Safes Framework, public good, RDS, Researcher Access Service, Private Sector Access Framework). • Explain that some of the scenarios we will be discussing today include some sensitive information or may cover situations that affect you or people you know. Show the wellbeing signposting slide and explain that these are some support links in case any of the discussion today raises things that are upsetting. Also reassure participants that it is fine to step away from the discussion if you need a breather, just let your facilitator know.

Time allocated	Discussion structure	Objectives	Questions and materials
18.10 Breakouts			
18.10-18.15	Introductions Breakout 1 5 mins	Warm up and checking in	**Introductions (5 mins)** Facilitator introduces themselves and thanks for participating. Collects permission to audio record the discussion and then asks participants to introduce themselves. • What has stuck in your mind most since the first session? • We shared all the post-its from the first workshop before the session where everyone wrote down what the idea of "Public Good" means to them. Did you get a chance to look at that again, and what did you think of the answers people gave? Was there anything you read that changed your own view? (facilitator to share screen if needed) • Have any new questions or issues come up for you since the last session? • How do you feel about the plan for today?
18.15 Plenary			
18.15 - 18.30	Presentation to inform participant learning 15 mins	Further information/context setting	**Plenary presentation 1 (live) (15 mins) - The Researcher Access Service and public good** • Brief recap on who RDS is • Introduce Researcher Access Service – what it is, how it came about, why it's needed, who is involved in the RAS panel • How public engagement has informed the RAS to date • Introduce Public Impact Advisory Group pilot ○ Background (i.e. why/how the pilot came about) ○ What it will look like ○ When/how it could be involved in decision-making ○ Recapping how this group will inform the next stage of development ○ Broaden access to different types of data, not just health ○ Inform the decision making of RAS Panel ○ Inform set up of PIAG • Walk through of a health project example and the public good criteria that is currently being used for access to health data for research, how it was assessed, and how it is tracked/reported on)
Breakouts			

Time allocated	Discussion structure	Objectives	Questions and materials
18.30 – 18.40	Reflections 10 mins	To enable participants to reflect and gather initial views, and form any questions	**Reflections on presentation 1 (5 mins)** • What are your reflections on what you've just heard? • Was there anything that surprised you or stood out to you? **Question forming (5 mins)** • What questions/comments do you want to ask/make to [speaker]? • What is our priority question? And what about a back-up in case our question is asked by another group? • Does anyone want to ask our question on behalf of the group?
18.40 – 18.45	**Q&A** **5 mins**		Plenary (5 mins): chair to facilitate Q&A session, with facilitators asking the questions or calling on participants to ask their question.
Breakouts			
18.45 – 19.20	**Group discussions and deliberation on RAS and public involvement** **35 mins**	Scenario-based discussion to start addressing Q1 (How can we tell when the use of data about people in Scotland for research is in the 'public good', and when it is not?) Discussion to start addressing Q2 (How should the public be involved in decisions about	**Scenarios (35 mins)** We're now going to look at a scenario… Facilitator load miroboard and share screen, reading through research summary first. Check participants are clear about it / clarify any questions as needed then read through scenario. Scenario: Group 1 (Facilitator 1): A Group 2 (Facilitator 2): B Group 3 (Facilitator 3): C Facilitator to highlight who is requesting access, what type of data they need, and how long they need the data for, and why need it. Then ask: • Who could benefit from it? IF NEEDED, PROBE ON: ○ Which individuals? How? ○ Which groups? How? ○ Or all of Scotland? How? ○ What impact(s) could it have? How big would that impact/those impacts be? • What risks are there? IF NEEDED, PROBE ON:

Time allocated	Discussion structure	Objectives	Questions and materials
		sharing data about people in Scotland?)	o Stigmatising certain groups? o Unfair treatment of/decisions about certain groups? o People's privacy? o Bias? o Representativeness and accuracy of decisions? o Financial? o Legal? o How big would that risk/those risks be? o Who could be affected if these risks became reality? How would they be affected? • What other information do we need to determine whether or not this project is for a public good? Why is that important? • If RDS was to provide this type of data to organisations like this one, do you think the benefits outweigh the risks? Why/why not? Facilitator note down benefits and risks on post its as mentioned by participants. Now if we assume that this project had been approved and has gone ahead... • How would we check that the project had done what it said it would? o What, if any, tracking or monitoring should be in place? o Currently, all approved projects are published on the RDS website. RDS are considering asking organisations to report back to them after a certain amount of time or at the end of the project. What are your thoughts on that approach? ▪ What kind of things would you want to know about? ▪ How much time after should they report back? (if response is "it depends, probe on what it would depend on in this scenario) Facilitator note down benefits and drawbacks of public involvement on post its as mentioned by participants. [IF TIME] Now thinking about public involvement… Facilitator recap on the PIAG (Public Impact Advisory Group which RDS would like to set up to involve members of the public in discussions about applications before they go to the approvals panel), then ask: • What would the benefits be of involving the public in decisions about this project? • What would the drawbacks be? • On balance. do you think the Public Impact Advisory Group should be involved in decisions about this project? o If no, why not?

Time allocated	Discussion structure	Objectives	Questions and materials
			○ If yes, why? • Is there any other information that you think members of the Advisory Group would need to help them review a project application? **Facilitator sends participants on 10 min break by 19.20**
BREAK (10 MINS): 19.20-19.30			
Plenary			
19:30 – 19.45	**Presentation to inform participant learning** **15 mins**	Further information/context setting	**Plenary presentation 2 (live) (15 mins) - Private Sector Access Framework** • Background – why is an approach to private sector access needed • RDS position on private sector access • Introduce PSAF – what it is and work to date • Introduce models of private sector involvement • Give overview of opportunities and risks with private sector involvement • Introduce concept of benefit sharing and how this relates to the models • Reiterate role of group (i.e. to consider these models and share views on what is acceptable/unacceptable in relation to private sector access to data)
Breakouts			
19.45 – 19.55	Reflections 10 mins	To enable participants to reflect and gather initial views, and form any questions	**Reflections on presentation 2 (5 mins)** • What are your reflections on what you've just heard? • Was there anything that surprised you or stood out to you? **Question forming (5 mins)** • What questions/comments do you want to ask/make to [speaker]? • What is our priority question? And what about a back-up in case our question is asked by another group? • Does anyone want to ask our question on behalf of the group?
Plenary			

Time allocated	Discussion structure	Objectives	Questions and materials
19.55 – 20.00	Q&A 5 mins	To allow participants to pose questions to speakers	Plenary (5 mins): chair to facilitate Q&A session, with facilitators asking the questions or calling on participants to ask their question. Chair to remind participants that experts can be called over to tables at any points throughout the day to answer questions.
BREAK (10 MINS): 20.00-20.10			
Breakouts			
20.10 – 20.40	**Group discussions and deliberation on private sector access public involvement** **30 mins**	Scenario-based discussion to start addressing Q3 (When is it acceptable and when it is unacceptable for private sector organisations to access data about people in Scotland?) Discussion to start addressing Q2 (How should the public be involved in decisions about sharing data about people in Scotland?)	**Scenarios (30 mins)** We're now going to look at a scenario and consider the three models for private sector access to data… Facilitator load miroboard and share screen, reading through task sheet first. Check participants are clear about it / clarify any questions as needed then read through scenario. Scenario: Group 1 (Facilitator 1) – F Group 2 (Facilitator 2) – D Group 3 (Facilitator 3) - E Facilitator to highlight who is requesting access, what type of data they need, and how long they need the data for, and why need it. Then ask: • What impact(s) could it have? • Who could benefit from it? IF NEEDED, PROBE ON: o Which individuals? How? o Which groups? How? o Or all of Scotland? How? o What impact(s) could it have? How big would that impact/those impacts be? • What risks are there? o Stigmatising certain groups? o Unfair treatment of/decisions about certain groups? o People's privacy? o Bias? o Representativeness and accuracy of decisions? o Financial? o Legal? o How big would this risk/those risks be? o Who could be affected if these risks became reality? How would they be affected?

Time allocated	Discussion structure	Objectives	Questions and materials
			Now let's look at the ways in which this company could access the data. Model 1: Research for, but not by, a private sector company Model 2: Research by a private sector company working with others in the public sector Model 3: Research by a private sector organisation not in partnership Model order: Group 1 (Facilitator 1) – 3,1,2 Group 2 (Facilitator 2) – 2,3,1 Group 3 (Facilitator 3) - 1,2,3 Ask for each model: • What are your thoughts on this model? ○ What do you like about it? ○ What do you dislike about it? • To what extent, if at all, would this model ensure the benefits we just discussed? Probe fully on how/reasons why not • To what extent, if at all, would this model address the risks we just discussed? Probe fully on how/reasons why not • **Would it be acceptable or unacceptable to you if RDS was to allow [organisation] access to this data under this model? Why?** ○ **[If unacceptable] What would need to be in place to make it acceptable?** • If the research went ahead under this model, how would RDS make sure that the project is doing what it said it would? ○ Are these the same or different to what we discussed in the early scenario? Probe on reasons for any differences [IF TIME] After discussing all models, ask: • Thinking about this scenario, which model do you think would work best? Why? • And which model do you think would work least? Why? [IF TIME] And now coming back to public involvement… Facilitator recap on the PIAG (Public Impact Advisory Group which RDS would like to set up to involve members of the public in discussions about applications before they go to the approvals panel and refer back to benefits/drawbacks previous discussed:

Time allocated	Discussion structure	Objectives	Questions and materials
			Facilitator to send participants on 5 min break if group needs it between this and final session (wrap up can be reduced to 1 min).
Stay in breakouts			
20.40 – 20.55	**Initial conclusions** **15 mins**	Summarising initial responses to key overarching questions	**Initial conclusions (15 mins)** In this last section we're going to bring together some initial conclusions to those three key questions that we showed you at the beginning. Question order: Group 1 – 1, 2, 3 Group 2 – 3, 2, 1 4. **How can we tell when the use of de-identified data about people in Scotland for research is for the 'public good', and when it is not?** Facilitator refer back to notes on first scenario and ask if there's anything to add, change or expand on (asking participants to explain their reasons and asking others in the group whether they agree or disagree – reassuring that we don't have to reach a consensus, but that we want to make sure that everyone has had their say). If completed, encourage participants to revisit the benefit/risk quadrant. Additional probes if needed: • Thinking back on the scenarios we looked at earlier today (briefly recap/summarise), what other issues did they raise that might be relevant to this question? • [The scenario] was placed [summarise placement of scenario on chart] – why do you think that was the case? What does that mean for determining whether or not a project is for a 'public good'? • [IF TIME] What if this data access request didn't come from [public sector organisation], but a private sector organisation instead? Would that change things or not? Why not/how so? 5. **How should the public be involved in decisions about sharing de-identified data about people in Scotland?** Facilitator refer back to flipchart/notes and ask if there's anything to add, change or expand on (asking participants to explain their reasons and asking others in the group whether they agree or disagree – reassuring that we don't have to reach a consensus, but that we want to make sure that everyone has had their say)

Time allocated	Discussion structure	Objectives	Questions and materials
			• Should public involvement look the same or different for different types of research project? Why/why not? • Should public involvement look the same or different if the private sector was involved in a research project? Why/why not? 6. **Which methods for private sector organisations accessing data about people in Scotland are acceptable?** Facilitator refer back to flipchart/notes and ask if there's anything to add, change or expand on (asking participants to explain their reasons and asking others in the group whether they agree or disagree – reassuring that we don't have to reach a consensus, but that we want to make sure that everyone has had their say) Additional probes if needed: • Thinking back on the scenarios/models we looked at earlier today (briefly recap/summarise), what other issues did they raise that might be relevant to this question?
Plenary			
20.55 – 21.00	**Wrap-up** **5 mins**		Chair to close the day: • Brief overview of what has been covered • Brief overview of next steps i.e. similar workshop with in-person group; RDS will reflect on the feedback so far and share some draft conclusions at the final online workshop • Chair highlights short survey participants asked to complete after this workshop • Reminder of incentives process **Thank participants and close** Facilitators and support will only leave once all participants have left. Facilitators to make a note of any unanswered questions to share with research team.
			<u>Short survey questions for after the workshop</u>

RDS deliberative

Workshop 3: online deliberation, Tuesday 20 May 6pm-9pm

Group of 35 participants, with 5 pre-assigned groups (of 7 participants each)

Overarching objectives:

- Participants hear playback of key themes/findings emerging from first two workshops and RDS's response / any updates on Researcher Access Service, Public Impact Advisory Group and Private Sector Access Framework.
- Participants review and ratify draft conclusions on overarching questions.
- Participants reflect on process and engagement with key terms.

Time allocated	Discussion structure	Objectives	Process, Questions and materials
17.30 – 17.50	**Set-up: Facilitators check in** **20 mins**	Ensure technology is set-up correctly before participants enter the room	**Chair, facilitators and tech team only** • Test link, mics and cameras. • Test who has the host/co-host function and ensure it is allocated to the right team member(s) for assigning break out rooms. • Make all facilitators co-hosts. • Change screen name to NAME – Org – Chair/Facilitator. • Check everyone is on the WhatsApp group for facilitation team to be able to ask questions • Facilitators, speakers and observers allowed in early and asked to change screen name • Meanwhile tech support is assigning participants who are in the waiting room, facilitators, experts and observers to break-out rooms.
17.50 -18.00	**Participant check-in** **10 mins**	Enable participants to get settled and resolve any tech issues	**Participants log into the online session** • Participants encouraged to join the Zoom session early to check in and check their video/mic. • Participants encouraged to get a pen and paper and get settled somewhere quiet. • Register as people join and change screen names as necessary to first name and first initial of surname (e.g. Jane S).

Time allocated	Discussion structure	Objectives	Process, Questions and materials
18.00 -18.10	**Welcome, and context setting** **10 mins**	To welcome participants and set out plan for final workshop and create sense of closure, achievement	Participants allocated to break-out groups, but not put in them **Ipsos Chair to welcome everyone to the final session:** • Chair gives everyone a warm welcome and thanks them for returning for the final workshop. • Chair to remind participants of the overall aims and purpose of the workshops (including reminder of who Ipsos and RDS are) and the role that participants are playing. • Recap who is here, including introducing any new speakers/observers who were not present at previous workshops. • Recap on key themes from previous discussions before recapping on key Qs: ○ **How can we tell when the use of de-identified data about people in Scotland for research is in the 'public good', and when it is not?** ○ **How should the public be involved in decisions about sharing de-identified data about people in Scotland?** ○ **Which methods for private sector organisations accessing data about people in Scotland are acceptable?** • Run though the agenda and what to expect from final workshop (i.e. look at public involvement – PIAG – before presenting draft conclusions, updates from RDS, reviewing and ratifying conclusions, and some final reflective discussion). • Explain confidentiality and anonymity, housekeeping, ground rules.
18.10 Breakouts			

Time allocated	Discussion structure	Objectives	Process, Questions and materials
18.10 – 18.15	Introductions Breakout 1 5 mins	Warm up and checking in	Introductions (5 mins) Facilitator introduces themselves and thanks for participating. Collects permission to audio record the discussion and then asks participants to introduce themselves. • What stood out most to you from the summary Chair just shared? ○ Has anything in particular stuck in your mind since the last session? • How do you feel about the plan for today?
18.15 Plenary			
18.15 - 18.25	**Presentation 1** 10 mins		**Chair to introduce presenter to talk about the Public Impact Advisory Group, which is about addressing the overarching question: How should the public be involved in decisions about sharing de-identified data about people in Scotland?** Chair to highlight that some participants will have discussed the role of the PIAG to some extent already, while others won't have yet. So the next discussion will be an opportunity for everyone to share views on the PIAG before we move onto conclusions. Chair will explain that we won't have a Q&A session, but if anyone has any questions to let their facilitator know so that a representative from RDS can be brought into your group to answer it. **Plenary presentation 1 (live) (10 mins) –** Public Impact Advisory Group • Recap on what has been covered so far • How the PIAG will be recruited • How the PIAG will fit into the researcher access approval process • Current thinking around the PIAG's level of involvement
18.25 Breakouts			
18.25 - 18.45	**Reflections on presentation 1 and discussion on PIAG** 20 mins		**Reflections on PIAG** Facilitator explains that we will now have a discussion about the Public Impact Advisory Group and its potential role. Let participants know that if they have any questions about the PIAG, we can call in a representative from RDS to answer it [facilitator to use judgement as to whether a question requires clarity from RDS, or whether it is a point for discussion].

Time allocated	Discussion structure	Objectives	Process, Questions and materials
			Recap if needed: the Public Impact Advisory Group is something RDS would like to set up to involve members of the public in discussions about applications before they go to the approvals panel. – **Based on what you've heard so far, what do you think the potential benefits of having a Public Impact Advisory Group are?** – **What are the potential drawbacks?** – **The current thinking is that the Public Impact Advisory Group would be involved in some applications (not all of them). What are your thoughts on this approach?** ○ How should RDS decide which cases to bring to the Public Impact Advisory Group? Possible probes: ▪ Based on the level of risk? If so – what kinds of risks?* ▪ Based on the type of data being requested? If so – what types of data should the PIAG be reviewing? ▪ Based on the type of organisation? If so – which types of organisations should the PIAG be reviewing? – *If helpful, facilitator to show summary of scenarios covered in previous sessions and the risk/benefit chart showing how those who took part in the in-person workshop weighed up the risks and benefits. **Summary of risk/benefit chart:** Scenario A (children in kinship care), C (income and education link) and D (Educ8 - tutoring) were generally thought to be high risk and high benefit Scenario B (youth justice system) was generally felt to be low risk, low benefit Scenario E (Nu Homes – housing developer) and F (EcoEnergy – home energy improvements) were generally felt to be high risk, low benefit - [Facilitator show researcher access request process diagram] **For the applications that are selected, the current thinking is that the Public Impact Advisory Group would discuss them before the approvals panel. What are your thoughts on this approach?** If other stages of involvement mentioned: Why do you think it is important that the Public Impact Advisory Group be involved at this stage too? ▪ What are the benefits? ▪ What are the challenges? How might these be overcome?

Time allocated	Discussion structure	Objectives	Process, Questions and materials
			Facilitator to note down key themes for recapping on when it comes to ratifying draft conclusions on public involvement.
18.45 Plenary			
18.45 - 18.50	Feedback 5 mins		**Brief feedback (5 mins)** Chair to invite facilitators to briefly summarise feedback in relation to the PIAG, focusing on how RDS should decide which applications should go to the PIAG.
18.50 - 19.05	**Presentation 2** 15 mins		**Plenary presentation 2 (live) (15 mins) – Draft conclusions** – Summary of answers to the three key questions from participant discussions. – Chair explains each group will review the draft conclusions and report back on any changes they've made. – Chair welcomes back presenters from RDS to provide a response to emerging conclusions (highlighting any developments or new thinking that has emerged based on the findings so far that RDS would like participants to consider as they finalise conclusions).
BREAK: 19.05 – 19.15 (Stay in plenary): at the end of break, chair explains plan for next session before opening breakouts.			
19.15 Move to breakouts			
19.15 to 20.15	**Groups discussion to ratify conclusions, including 10 min break**		Facilitators remind participants of the overall aims: to help inform decisions about how data about the Scottish public is accessed for research, and to answer the specific questions posed. Facilitator load miroboard and share screen showing draft conclusions. Facilitator to send participants on a 10 minute break after discussing the first or second question, depending on participant preference/energy levels. We will now review the draft conclusions and refine them based on everything we've heard and discussed. We will go through the draft conclusions together and you can tell me if there's anything you'd like to change, add or leave as it is. It's okay if you have different views on certain conclusions, I will make a note of that, and we will make sure that any differences in views are reflected in the report. We're more interested in making sure the conclusions reflect the full range of views, so you don't need to reach a consensus.

Time allocated	Discussion structure	Objectives	Process, Questions and materials
			Question order: Group 1(Facilitator 1) – 1, 2, 3 Group 2 (Facilitator 2) – 2, 3, 1 Group 3 (Facilitator 3) – 3, 1, 2 Group 4 (Facilitator 4) – 1, 3, 2 Group 5 (Facilitator 5) – 2, 1, 3 Facilitator to read through draft conclusions for each question, allow participants some thinking time, then go through the questions and live edit the conclusions as required (highlighting any direct edits to the conclusions, and noting different views on post-its where any disagreement arises). **Q1. How can we tell when the use of data about people in Scotland for research is in the 'public good', and when it is not?** • **Is there anything you would want to change?** Why is that? ○ Is there anything missing? What would you like to add? Why is that important? ○ Is there anything you're happy with as it is? What makes you say that? • **In RDS's response to these initial conclusions, we heard that research findings do sometimes lead to more questions than answers and that the public benefits may come about from further research, meaning the benefits from the original research are indirect (i.e. a stepping stone towards a change/improvement) rather than direct (i.e. leading directly to a change or improvement in something). What are your thoughts on that?** ○ Would research still be in the public good if it generates further questions? Why/why not? ○ What does this mean for our current statement that "use of data is not in the public good when the research objectives are vague, or the intended public benefits are unclear, hard to measure, and leave people with "more questions than answers" – is there anything we want to add or change here, or should it stay as it is? Probe fully for reasons • **[IF TIME] Let's think about different types of data (read out the reminder of data that RDS may provide access to in future). Are there any further**

Time allocated	Discussion structure	Objectives	Process, Questions and materials
			considerations for these different types of data when thinking about what is or isn't in the public good? ○ Facilitator to encourage participants to consider the scenarios they looked at last time and show scenario summary slide if helpful. social services records, education records, NHS records, property information, landownership, employment status, SIMD (characteristics of an area), justice records, benefits records, household income. ○ Do we need to make any edits to our conclusions, or the examples, to reflect these considerations? • **[IF TIME] Let's imagine that all these conclusions became the criteria that was provided to a researcher. How many criteria would you expect the researcher to meet in their application?** Probe fully for reasoning ○ Would it be acceptable if the researcher's application prioritised one criterion over another? • **[ALLOW TIME BEFORE MOVING ON] Of everything here, what is most important to you? What makes you say that?** ○ If important things mentioned are all in green box, ask: And looking at the statements in the red box, is there anything there that feels particularly important? Why is that? ○ If important things mentioned are all in red box, ask: And looking at the statements in the green box, is there anything there that feels particularly important? Why is that? **Q2. How should the public be involved in decisions about sharing data about people in Scotland?** (feedback on Public Impact Advisory Group as part of the Researcher Access Service): facilitator reiterate that these emerging conclusions were based on views shared in the in-person session, but may need to be updated based on further discussions this evening. • **Based on what you've heard and discussed this evening, is there anything you would want to change?** Why is that? ○ Is there anything missing? What would you like to add? Why is that important? ○ Is there anything you're happy with as it is? What makes you say that?

<table>
<tr><th>Time allocated</th><th>Discussion structure</th><th>Objectives</th><th>Process, Questions and materials</th></tr>
<tr><td></td><td></td><td></td><td>Earlier we discussed the benefits and drawbacks of involving the public in the application process. Facilitator recap on views shared. Do we need to edit these conclusions to reflect this?
We also discussed what types of applications the Public Impact Advisory Group should be involved in. Facilitator recap on views shared. Do we need to edit these conclusions to reflect this?
Why is this important?
[IF COVERED EARLIER AND THERE'S TIME] And we discussed what stages of the approval process the Public Impact Advisory Group should be involved in. Facilitator recap on views shared. Do we need to edit these conclusions to reflect this?
Why is this important?
<p>Q3. When is it acceptable and when it is unacceptable for private sector organisations to access data about people in Scotland? (feedback on the 'models')</p>
Is there anything you would want to change? Why is that?
Is there anything missing? What would you like to add? Why is that important?
Is there anything you're happy with as it is? What makes you say that?
In RDS's response to these initial conclusions, we heard that private companies would be required to publish a summary of the research and details of the data used and a public good statement. But the requirement of private companies to publish the findings from the research may put companies off from doing the research in the first place (e.g. due to potential commercial sensitivities). What are your thoughts on that?
Should the requirement to publish findings be negotiable, or non-negotiable? What makes you say that?
If negotiable – what should private companies be expected to make publicly available at a minimum?
What does this mean for our current statement under model 1 – "there should be a requirement on the private company to make research findings publicly available to avoid the risk of them 'cherry-picking' the findings that align with their commercial motives" – should it stay as it is or should we revise it? Probe fully for reasons</td></tr>
</table>

Time allocated	Discussion structure	Objectives	Process, Questions and materials
			• **[IF TIME] In RDS's response, we also heard that these models could be more fluid in reality e.g. an organisation might exist both in the private and public sectors, or a private sector organisation might have a long-term relationship with a public sector organisation like the NHS) – what are your thoughts on this?** ○ Does this make a difference to the considerations we have written down here? ○ Would having more than one private organisation involved make a difference? If no – why not / if yes – how so/under which models? ○ What does this mean for our conclusions? Is there anything we want to add or change? • **[IF TIME] Our draft conclusions talk about safeguards being in place for private sector access to data (read relevant statement under model 1 and model 3) - what sort of safeguards should be in place for private sector access to be acceptable?** ○ Does this apply to one model in particular, or across all the models? ○ Does this apply to all types of data being accessed, or specific types? ○ Do we need to add or change any of our conclusions to reflect this? • **[IF TIME] Under model one, the draft conclusions also talk about vetting a company (read relevant statement) – what would be the key things to check?** • **[ALLOW TIME BEFORE MOVING ON] Considering everything we've discussed, what is the most important thing that you feel RDS should consider for any future private sector access to data?**
20.15 Plenary			
20.15 to 20.30	**Feedback in plenary**		**Feedback (15 mins)** Chair to invite facilitators to share screen and briefly summarise key edits to conclusions and reasons for this.
20.30 Breakouts			

Time allocated	Discussion structure	Objectives	Process, Questions and materials
20.30 - 20.55	**Group discussions** **25 mins**		**Reflections on emerging themes/conclusions from other groups (5-10 mins)** • **What stood out to you from the other group's edits?** ○ Was there a change or lack of change that you really support? What makes you say that? ○ Was there a change or lack of change that you really don't support, or disagree with? What makes you say that? • **Was there anything raised by another group that has changed your thinking on our edits?** ○ Is there something you'd like to change now? Why is that? **Reflections on concepts and informing the public (15-20 mins)** Thank you very much for your work on these conclusions. We're almost at the end of the session now. In these workshops, we've learnt about and discussed concepts that you might not have considered before, from what we mean by data about people living in Scotland, how it's used for research, how it's kept secure, through to how RDS and similar organisations provide access to data for research, the types of data that could be accessed and the organisations that might get access, and the importance of research being in the public good. To finish off, we're interested to hear your thoughts on how these concepts could be communicated with the general public, based on your experiences of taking part in this research. • **Thinking back, which topics, if any, do you think have been most interesting to learn about?** Ask openly first, but if discussion slow could probe on: ○ How data about people living in Scotland is protected ○ The concept of 'public good' when using data about people living in Scotland for research ○ RDS and its role ○ The Researcher Access Service ○ Public Impact Advisory Group ○ The models for private sector access to data ○ For all issues/topics mentioned, ask: ▪ Was there anything in particular that you heard or saw that made this topic/issue <u>more engaging</u>? ▪ Was there anything in particular that you heard or saw that made this topic/issue <u>less engaging</u>?

Time allocated	Discussion structure	Objectives	Process, Questions and materials
			▪ Probe fully on things participants remember that helped or not, and reasons why • **And which topics, if any, do you think have been the most challenging topics to learn about?** Ask openly first, but if discussion slow could probe on: ○ How data about people living in Scotland is protected ○ The concept of 'public good' when using data about people living in Scotland for research ○ RDS and its role ○ The Researcher Access Service ○ Public Impact Advisory Group ○ The models for private sector access to data ○ For all issues/topics mentioned, ask: ▪ Was there anything in particular that you heard or saw that made it easier to learn about this topic/issue? ▪ Was there anything in particular that you heard or saw that made it harder to learn about this topic/issue? ▪ Probe fully on things participants remember that helped or not, and reasons why • **Now thinking about everything we've learned about in these workshops, what topics/issues (if any) do you think are most important for the general public to know about?** ○ Why is it important for the general public to know about this? ○ Do you have any thoughts on how RDS should inform members of the public about [topics/issues mentioned]? ▪ Where would you personally look for this type of information? ○ Anything else? Repeat probes for all topics/issues mentioned **IF TIME:** With the few minutes remaining, I'd be interested to hear your thoughts on this process and your participation… • What, if anything, have you enjoyed most about being part of these discussions? • What, if anything, have you not enjoyed as much? • What, if anything, has been the most challenging part? • What, if anything, will you take away from the process?

Time allocated	Discussion structure	Objectives	Process, Questions and materials
			• What advice can you give us or RDS to make these kinds of processes better in future?
20.55 Plenary			
20.55 – 21.00	**Recap, next steps and close**		**Ipsos Chair to close the day (5 mins):** • Recap on what has been discussed • Chair to introduce [speaker from RDS] to explain how the findings will be used / publicised. • Chair to thank participants for their engagement throughout the workshops • Chair to explain final steps: they will receive their final payment from FieldMouse and receive a wrap up email from Ipsos (which will include a consent question about sharing email address with RDS so that they can be kept informed). **Thank participants and close** Facilitators and support will only leave once all participants have left.

D. Q&A document

Question	Answer
Public good	
Who actually decides if is something is in the public good – i.e. what body, and what governance of that body in in place?	There is a legal definition from the UK Statistics Authority as part of the Digital Economy Act. Each service providing access to data has appointed individuals, called data access committees or panels, to make that decision. All are slightly different but are usually made up of people with responsibility for the data, data privacy professionals, researchers advocates and members of the public. For the RDS service, whether a project meets the public good is currently determined by our Researcher Access Service approval panel – made up of representatives from Research Data Scotland and Public Health Scotland. However, in future, we hope to have a public impact advisory group who will provide their feedback on whether they feel the proposed research is for the public good. You can find out more about how the public good is currently reviewed here.
There was mention of long-term and short-term benefits – how are these time periods defined?	These are defined on a case-by-case basis, of course some research during the COVID pandemic for example influenced some immediate policy changes, but for most research it is understood that most benefits will not appear for at least a year after researchers start analysis, but may take longer. In the context of judging public good, we should consider the whole picture and ensure there could be no risk to individuals or society further down the line

Does the definition of public good change depending on government priorities / the government of the day? **Could the political environment impact on the rules around how data is managed and used? For example, if a new government is voted in in 10 years, could they change the approach? Also, if public services such as the NHS become privatised, would that mean a private company then has our data? Are there any long-term safeguards to stop politics impacting how data is managed?**	Organisations involved are a mix of government, public sector and charity. There are legal expectations such as the Digital Economy Act or Health Act Scotland about making data available for research. Organisations like the Office for Statistics Regulation have done lots of work to define public good and research needs around it which maintain consistency through governmental changes. Whilst priorities of a government can always affect things in a broad sense, there would be a lot of other factors, organisations, and people independent of government contributing to any change.
How can it really be based on a social contract if most people don't know about all this?	Public sector organisations that collect and process personal information inform the public about how their data is used through privacy notices on their websites and often through information displayed on their public premises. RDS do not collect data but instead facilitate access to existing data collected by other public sector organisations. We continue to engage the public on the work we do (which is why we're talking with you now) as well as using findings from other public consultations but recognise there is more to do. We have funded/supported more work in this area through our public engagement fund and attendance at public events, such as science festivals. We also put as much information about our work as we can on our website in a clear way. This work relies on public trust, and we will continue to work with the public to ensure our work has the confidence and support of the public. We will be covering how we can communicate our work better in future workshops.

How do you measure public good? What's the benchmark?	We publish information on our website about how we currently use public benefit definitions to assess projects in the Researcher Access Service. We are engaging the public again to ensure our definitions command the confidence of the public while we further develop the service, because we recognise that public engagement is not a one-off exercise.
I worry about inclusivity. What's the cut-off for ensuring public good? Does it have to affect a particular amount of people to be approved?	No, there is no set amount of people that have to benefit from the work for a project to be approved. The benefits just have to outweigh the risks. A good example is studies that consider rare health conditions. It may only affect a small number of the population, but it would still be approved as the benefit is higher than any potential risks to data access.
Public sector/Private sector	
Can research be in the public good if it is being driven by interests of profit? Is the test the same as for public sector researchers?	Whilst a company may be driven by profit, this cannot be their primary reason for requesting access to data from a Trusted Research Environment. The work outlined must be for the public good. Research can be done in the public good if driven by interests beyond profit, even if there is a secondary goal of continued profit for a company. One example, that was recently reviewed by Research Data Scotland's public panel, was for an organisation to access 1.6 million brain scans to help predict a person's risk of dementia. Whilst initially this is being researched in the public sector, this could be used in the future (subject to further approvals) by a private company to ensure the project continues and appropriate medicines could be developed. As Research Data Scotland does not currently allow access to data by private sector companies, the access and review process has not yet been agreed, and this is something we look forward to discussing with you at the next workshop.
How do we know what private companies will do with the data once they have it ('once the data is out of the bottle')? E.g. what control is there over what countries they send	International access is judged on a case-by-case basis All data which an organisation has access to is de-identified, and it can only be accessed by logging in to our secure locked down National Safe Haven through a secure internet connection. Researchers don't have the ability to put anything in or take anything out.

it to? (We have GDPR but not all countries do).	Once researchers have analysed it in a safe environment then they are only allowed to take out summary outputs, which are also checked to make sure no one can be re-identified. All statistics researchers produce are checked by National Safe Haven support staff to make sure these match the purpose of the approved project and do not identify people in the data. GDPR only applies to identifiable data, we are working with de-identified data.
What is the definition of public sector and what is the interaction with the voluntary sector (which hasn't been mentioned)?	Data: Public sector data is information collected throughout a person's lifetime. Public sector bodies are the organisations that collect this, such as the NHS, local authorities, government bodies, schools, social care organisations, and more. For example, GP records might contain data about an individual's medical history and diagnoses. Similarly, local authorities and government bodies may store information about a person's employment status and housing information. Most public sector organisations will also hold general person-level data such as an individual's name, date of birth and address. This information is known as public sector data or administrative data. We do not currently hold information from the voluntary sector. Access: Access is currently limited to researchers from academic institutions, public sector or voluntary sector. You can find further information available on our website.
Who sets the standards for private sector access and what happens if the private sector does something wrong?	There is currently no universal standard in place for private sector access to public sector data. Part of our discussions with you at the next workshop will be to help develop and define how standards are set at Research Data Scotland, which aims to create one source for data access. If the private sector does something wrong, they would be considered in breach of their user agreement. A breach of user agreements for researchers is currently a criminal offence with a penalty of 2 years prison sentence and £11k fine. It is envisioned that private

	sector researchers breaching user agreements, at minimum, would be subject to this same legal prosecution. However, the fine is likely be increased for private sector organisations.
How is the output assessed? Can private sector companies be followed up with after?	Research Data Scotland don't currently allow access to data by private sector organisations. Private sector organisations can be followed up with after their research is complete to ensure the appropriate outputs are in place. Private sector organisations will also be subject to user agreements they have signed and agreed with Research Data Scotland. However, these agreements are not yet in place and at the next session we will be looking for your input into the development of our framework. We will welcome your feedback in this area.
RDS	
Who is funding RDS? // How much funding has the government put into RDS?	Research Data Scotland receives funding from the Scottish Government and also research grant funding. You can find full details of how we are funded on our website as well as our financial records in Governance documents.
What does the RDS charity set-up look like? Does the charity commission have a say on the role of RDS / a governance role (seeing as it is a charity)?	Research Data Scotland is a not-for-profit charity regulated by the Office of the Scottish Charity Regulator (SC051305). Setting up RDS as a charitable organisation provides some nimbleness to our work, while also maintaining charitable aims and working with a range of partners, and allows funding to be carried over year on year and hence greater stability over time to lead the system transformation needed. RDS is also governed by our Board of Directors and subcommittees. You can find out more about our governance here: Governance | Research Data Scotland.

RDS seem to be the key player in this, could anyone set up this sort of data system or is it government-determined that RDS can be the only ones? i.e. is there a risk that someone else could set a similar system up with our data, as a private company, and not be as scrupulous?	Other organisations exist across the UK that do similar work for particular types of data. In theory, another organisation could be created to do something similar in Scotland. However, providing a system to access de-identified datasets requires strict access controls agreed with the organisations that own the data, such as the NHS, National Records Scotland and others. A lot of partnership working and agreements have to be in place and trust built. Therefore, it is unlikely that an unscrupulous organisation would be able to establish something like Research Data Scotland.
Why is RDS doing this research NOW (as opposed to a couple of years ago, for example)?	Research Data Scotland has been established since 2021. During this time, we have worked closely with others in the sector to be informed by other public engagement work carried out across the UK. We have also done some of our own work on a smaller scale and with other organisations. For example, we have a public panel which meets five times a year to help us. A lot of work over the last few years has been in preparation for this next step in the Researcher Access Service. We are now at a point where we are making a big step forward and putting in place systems that we want the public to directly input into.
Who is asking RDS to do this research?	RDS is carrying out this work as part of their commitment to involving the public in our work. We work closely with lots of different organisations which also carry out public engagement such as this to share findings and make sure we are having relevant conversations with people.
I know that the presenter mentioned that RDS charges for the service and not for the data itself, but how do they exactly structure that pricing model and how do they make sure that that mitigates against the risk of it being a pay for play type model where bigger companies have more access to data or more access to the	The pricing model has been set up to cover costs for - Research Data Scotland and eDRIS (a public sector partner) - Technical infrastructure - Indexing (organising the data for researchers) Researchers are charged based on the size and complexity of their projects. They usually include these costs in their grant applications in the same way as other costs they would need to complete their project.

service than smaller companies? And how do they balance that with the research for the good side of it too?	Applications by researchers to access data are reviewed separately from the tea that calculate costing. Each project is judged by the public good it can do and whether they meet the five safe requirements. This assessment of public good is carried out by an independent panel. We are working to be more transparent around panel decision making building on the publicly available register of approved projects on our website.
It would be interesting to know how many employees there are and what bands of employees there are. There must be some who are technical and some who are administrative and some who are, I don't know, just executive?	Research Data Scotland has 36 employees with a mix of expertise and roles. We have a: - Data team who support public sector organisations to make their data available for research -Digital team who build and manage our Researcher Access Service and website -Partnership and Comms team who make sure we are working effectively with the whole sector whilst being transparent about what we do and engaging with public and users. - Operations Team who manage the finances and HR and general running of RDS A list of our team members can be found on our website You can also find all of our staffing information, accounts and annual reports publicly online.
Who are on these committees who decide on Access? Are they appointed by the government? Are they appointed by RDS? Can I apply? Can I get. You know, it's a bit vague. Let's say who actually makes the decisions on Access?	The Research Access Service approval panel – made up of representatives from Research Data Scotland and Public Health Scotland – will make the final decision on who gains access. However, we soon hope to have a public panel in place – including representatives across Scotland to help provide feedback to applicants and inform Research Access Service approval panel decisions. We will be looking to gain your feedback on how this panel is involved in the review process.
A lot of the data in the census would be useful. Is there a tie up was between Research Data Scotland	Research Data Scotland are working with National Records Scotland and other public sector organisations to make their data available securely through the Researcher Access Service

and the National Records of Scotland?	
The presenter said that RDS's aim is to help researchers access data, but how do they gather the data in the first place? Because presumably there are lots of different organisations doing lots of different research, so how do they [RDS] get access to it?	Throughout a person's lifetime, data is collected about them. This includes records about the interactions people have with services including: • Records of births, marriages and deaths • Health records - hospital admissions or prescriptions • Education records - exclusions & absences, attainment, leaver destinations • Tax and social security records • Police, prison and court records • Children in care After removing identifying information, - such as names, phone numbers, and addresses - Research Data Scotland wants to bring more of this data together into one secure place which is under public sector ownership and management. That's because we're aiming to make it faster and simpler, whilst still secure, for researchers to access public sector data for research.
You mentioned that people apply to access the data. Does that mean RDS has all the data? You know, they can't just come to you and say, I want all the data based on, say, house sales. They must know that you have all this data to start off with, otherwise they couldn't request it. So there must be a catalogue of data that you hold that you then make publicly accessible?	The RDS metadata catalogue provides information about what datasets are available from various public sector organisations across Scotland. The information provided for each dataset includes top-level descriptions, dates of creation and updates, access details, and some additional technical information to help researchers understand whether the dataset will be useful for their research. Once researchers have identified one or more datasets that will be suitable for their study, they can begin the process of requesting access to the datasets.
Data/processes/security	
Is there a chance that a researcher could trace back to the source data, or somehow fill in the missing pieces	It would be pretty much impossible for a researcher to do this. When data is linked together using source data and then de-identified it is done by multiple different teams. These

to work out what data has been removed?	different teams do not even have access to each others workspaces. Only data that has then been de-identified is provided to the researcher.
While there are measures in places to keep data secure, you still hear about data breaches including in the NHS. What mechanisms are in place to protect against this? // What accountability mechanisms are in place for when there are data breaches? How can we be reassured that our data isn't being misused? Is there a government department that is accountable?	Here is a short video we produced to explain how data is kept secure Further information can be found here but here is a short excerpt: "The National Safe Haven meets several national and international security standards (ISO 27001:2013, Cyber Essentials, NHS England's Digital Security Protection Toolkit and is Digital Economy Act (2017) accredited) and penetration testing is performed on an annual basis. Penetration testing is where an external company is contracted to fully test the security of a system by simulating a cyberattack. This helps discover potential points that could be exploited and ensures the systems in place to deal with breaches operate accordingly." If a researcher is found to be breaking the user agreement they have signed then they can be fined £11k and receive a two year prison sentence. We will send a full response to this question next week.
How is the data checked for accuracy / how accurate is it?	The data is checked at multiple stages throughout the process. There is a first initial data quality check which scans all variables for high-level validity and highlight any potential inaccuracies and missing values as soon as the data is received. As the data is administrative or operational data there are limited opportunities to change the data that is received, and our default is to retain the data in the form that it is received. However, we do make informed decisions about the plausibility of values, for example if an age range was implausible. Once the personal data is removed and the data is being prepared to send to the National Safe Haven we carry out further checks to confirm that only data approved for research is being moved. There are further checks by partner organisations when they receive the data including the indexing report which evidences the linkage rate of the datasets (how

	many individuals were identified in Scotland's population spine), the partner organisations also check that they have only received the data that they have agreements to receive.
Do companies still buy data? Does it go to the highest bidder?	Research Data Scotland operates a cost-recovery model meaning we do not profit from researchers accessing data. Researchers are charged based on the size and complexity of their projects. Researchers usually include these costs in their grant applications in the same way as other costs they would need to complete their project. Applications by researchers from any company to access data are reviewed separately from the costing, and only fully qualified researchers will eventually be given access. Each application is judged by the public good it can do and that they are meeting the five safe requirements. There are lots of different types of data that exist. Research Data Scotland focuses on administrative data which is data already held securely about you from interactions with the NHS, education, police, census. It is only data which is de-identified.
Why not UK wide hubs?	Data is stored and coded in different ways across the UK making it hard to put all the data in one UK wide hub. There is a lot of work underway across the UK, which RDS is part of, to make it easier for researchers to access and compare data across the UK and to join up these different systems.
Isn't there more danger or a huge leak if all the data brought together?	Further information can be found here but here is a short excerpt: "The National Safe Haven meets several national and international security standards (ISO 27001:2013, Cyber Essentials, NHS England's Digital Security Protection Toolkit and is Digital Economy Act (2017) accredited) and penetration testing is performed on an annual basis. Penetration testing is where an external company is contracted to fully test the security of a system by simulating a cyberattack. This helps discover potential points that could be exploited and ensures the systems in place to deal with breaches operate accordingly."

	Security measures are very high and data leaks from Safe Havens are very rare. Even when the data is linked together it is all de-identified.
We heard a lot about use of the data, but what we didn't hear is where the data comes from to get into the system. So it would be good to hear more on that because obviously the outcomes depend on the quality of that data. So somebody must be deciding what data is?	Public sector data is information collected throughout a person's lifetime. Public sector bodies are the organisations that collect this, such as the NHS, local authorities, government bodies, schools, social care organisations, and more. For example, GP records might contain data about an individual's medical history and diagnoses. Similarly, local authorities and government bodies may store information about a person's employment status and housing information. Most public sector organisations will also hold general person-level data such as an individual's name, date of birth and address. This information is known as public sector data or administrative data. After removing identifying information, - such as names, phone numbers, and addresses - Research Data Scotland aims to bring more of this data together into one secure place which is under public sector ownership and management. That's because we're aiming to make it faster and simpler, whilst still secure, for researchers to access public sector data for research.
How does it actually work for a researcher, in practice, to use the safe haven? What does it involve for them?	Researchers must be from approved organisations and have up-to-date training from an approved training programme. Once a project is approved, researchers have to complete user agreements to receive their access instructions to the National Safe Haven. Researchers can only log into the National Safe Haven from an approved organisational computer and approved organisational IP address. Researchers receive a unique log in code every time they access. There are two security levels to pass each time you log in. When using the National Safe Haven system there are strict controls in place that make it impossible to access the internet and you cannot copy and paste or change file locations within your project area. You are working in a fully controlled space that only those approved on the project and eDRIS team members can access. When researchers want anything to come out of the National Safe Haven, they move it to a disclosure control folder and complete a form explaining the outputs they are requesting. These are checked by eDRIS staff before being released. Working in the National Safe

	Haven involves patience and adapting working style, as researchers do not have easy access to lots of functions we take for granted, like the internet or copying text across files. Find out more in our Trusted Research Environment Explainer video
How do people get our data for things like spam phone calls?	This is a different type of data collection and use than what RDS does. RDS would never hold information like telephone numbers or make them available to researchers. We only hold information that cannot identify you.
What are the sources of information in the safe haven?	Similar to Research Data Scotland, the Safe Havens access data from public sector bodies. Public sector bodies are the organisations that collect this, such as the NHS, local authorities, government bodies, schools, social care organisations, and more. For example, GP records might contain data about an individual's medical history and diagnoses. Similarly, local authorities and government bodies may store information about a person's employment status and housing information. Most public sector organisations will also hold general person-level data such as an individual's name, date of birth and address. This information is known as public sector data or administrative data. This information is not immediately accessible to any researcher, and access is given in only in a controlled environment, called a Trusted Research Environment. The researchers can only see the project folder they have been given access to.
Who is asking for the data, how does it come into the safe haven and where does it go after?	Data could be requested from any number of organisations, including universities, local authorities, government organisations and others. However, only qualified researchers will be able to see and use the de-identified data in a controlled environment, called a Trusted Research Environment. Whilst the safe haven controls access, the data remains in the Trusted Research Environment and never leaves this environment. Only the summary research findings can leave the safe haven (e.g. the 8% were identified as homeless example provided in our first workshop). These findings will be checked to ensure the data cannot identify individuals.

Does the information in the Safe Haven have an expiry? How long does it remain relevant?	The information in the Safe Haven is continually updated. For example, each time you visit the hospital it is recorded by the NHS. This information is then de-identified and made available in the Trusted Research Environment, though the information might be a couple of years behind.
What does it mean researchers can't take information out of a Safe Haven? What does it mean in practice for people working with the data?	Researchers must request outputs. They must complete a form every time they are requesting for outputs to be taken out of the National Safe Haven and only eDRIS employees can take outputs out and send them to researchers via email. Outputs are scrutinised by eDRIS staff and must never contain any information about individuals. Types of outputs that might be taken out can be graphs, figures, or summary tables of information but never individual values about single people.
They said birth date gets taken out in the process of de-identification, but imagine age must be important for researchers to know?	Full date of birth can be requested by researchers if it is critical to answering the research question though full date of birth is not available through the Researcher Access Service. Typically, month and year of birth is enough information for researchers to answer questions as this information is very important for lots of research questions. It is about the data minimisation principle - i.e. Researchers should only ask for and get to use the necessary data required to answer their research question.
How can we make sure that researchers use all the relevant data in their research and don't just select the data that backs up what they want to say?	Very good question. This is ensured at multiple points throughout the process. For example, at the enquiry stage, the Trusted Research Environment professionals will discuss and advise researchers on what data is most appropriate for their project. Checks and support like this will be ongoing throughout the research access process. In addition, anyone who accesses the data must be a qualified researcher and signs a user agreement. If they do not adhere to this agreement, they are in breach. If a breach occurs, both they and the organisation they represent will be penalised. For example, researchers can be fined £11k and receive a two-year prison sentence.
What protections are in place for the public? Is there a public forum that means we get to say no if we don't like a piece of research that could be or is being done?	All projects are published on the Data Use Register. Members of the public can contact Research Data Scotland directly about projects. Part of the next sessions will be looking at where members of the public sit in the process of approving and tracking projects.

How do people know what data they can access?	Research Data Scotland has worked with partner organisations across Scotland to build a single metadata catalogue that is available for the online. This is a catalogue which describes the datasets that are available for research purposes and outlines the organisation that has responsibility for the data.

E. Scenario descriptions

Scenario A: Understanding Experiences of Children in Kinship Care

A university researcher who specialises in children's welfare is studying what happens to children who are being raised by relatives or family friends instead of their birth parents (known as "kinship care"). The researcher is particularly interested in how this arrangement affects children's mental health and emotional wellbeing compared to other care arrangements.

The researcher wants to find out: How many children in kinship care experience mental health challenges? How does mental health affect their education and what they go on to do after leaving school? What interventions are most effective in supporting them?

What data does the researcher need?

Characteristics: age, gender, ethnicity

Social services records about children being cared for by relatives (collected by local councils and reported to the Scottish Government)

Mental health service usage and diagnosis records (collected by the NHS)

School grades and school destinations showing what these young people go on to do after leaving school (collected by schools and reported to the Scottish Government)

Remember: The researcher won't be talking directly with any children or families. Instead, they'll analyse existing information that has already been collected by government agencies. The data will be **de-identified** (names and personal details removed) but **linked** at an individual level, meaning the researcher can see one person's complete journey through social services, mental health support, and education without knowing who they are.

Scenario B: Youth Justice System Engagement and Future Outcomes

Community Justice Scotland is a public body conducting research into whether childhood interactions with the justice system is linked to criminal activity later in life. The organisation works closely with community organisations across Scotland and funds intervention programmes aimed at breaking cycles of reoffending.

Community Justice Scotland want to find out: Whether early involvement with the justice system (even as a witness) affects people later in life; how different factors (like education or support services) relate to young people committing crimes when they're older; and how to better support young people who have contact with the justice system.

What data does Community Justice Scotland need?

Youth justice records showing when young people had contact with police or courts (as witnesses, victims, or for offenses) before age 18 (collected by Police Scotland and courts)

Adult criminal justice records showing any convictions or interactions with the justice system after age 18 (collected by courts and Scottish Prison Service)

Education records (school grades, attendance, qualifications) and employment status (collected by schools and government agencies).

Remember: The researchers won't be talking directly with any young people. Instead, they'll analyse existing information that has already been collected by government agencies. The data will be **de-identified** (names and personal details removed) but **linked** at an individual level, meaning the researcher can see one person's complete journey through their interaction with the justice system and education without knowing who they are.

Scenario C: Understanding links between income, benefits and education

The Child Poverty Action Group (a charity that works to end child poverty) would like to understand if there is a relationship between receiving benefits and school attendance or success (e.g. good grades).

The Child Poverty Action Group wants to find out: Whether children from lower-income families face particular barriers to educational success; whether certain benefits help improve engagement with education; what happens after school for young people from different income backgrounds; and what support might help children from low-income families to succeed in education.

What data does Child Poverty Action Group need?

Household income and benefits information showing which families receive financial support and how much (collected by Department for Work and Pensions and local councils).

Characteristics of area (SIMD) and household (through Unique Property Reference Number), for example, the number of people in the household.

School attendance, grades and school leaver destinations, showing what young people go on to do after leaving school (collected by schools and Skills Development Scotland).

Remember: The researchers won't be talking directly with any families. Instead, they'll analyse existing information that has already been collected by government agencies. The data will be **de-identified** (names and personal details removed) but **linked** at an individual level, meaning the researcher can see one person's complete journey through their interaction with the benefit system and education without knowing who they actually are.

Scenario D: Educ8

Educ8 is a social enterprise startup founded by a student at the University of Dundee. The company provides affordable tutoring services to pupils from low-income households who are falling behind in school.

Educ8 wants to access data on attendance and attainment across Scotland's schools as well as household income data, to identify pilot areas for the tutoring service.

Educ8 wants to develop a "needs profile" that helps identify where their services would have the most impact, based on factors like: whether family income affects different subject areas differently; how attendance patterns relate to specific learning challenges and which combinations of factors create the highest risk of educational disadvantage.

What data does Educ8 need?

School attendance records, grades and school destinations showing what these young people go on to do after leaving school (collected by schools and universities).

Characteristics of area (SIMD) and household (through Unique Property Reference Number), for example, the number of people in the household; council tax bands,

Household income and benefit receipt information (collected by Department for Work and Pensions and local councils).

Remember: The researchers won't be talking directly with any students. Instead, they'll analyse existing information that has already been collected by government agencies. The data will be **de-identified** (names and personal details removed) but **linked** at an individual level, meaning the researcher can see one person's complete journey through their household characteristics, income and education without knowing who they are.

Scenario E: Nu Homes

Nu Homes is an international housing development company that wants to start building homes in Scotland. They want to build homes that meet local needs while being profitable. They want access to land data from Registers of Scotland, local authority development plans, and income and benefits data.

Nu Homes hope that access to this data will also help them identify areas of land that are likely to increase in value (e.g. due to population growth, household size, and increasing household incomes). They have the construction expertise to address housing shortages, but they need detailed data to ensure they build the right types of homes in the right places.

What data does Nu Homes need?

Local development plans and land availability.

Characteristics of area (SIMD) and household (through Unique Property Reference Number), including age of residents and number of residents.

Household characteristics linked with income and benefit information to identify the types of homes needed by low-income families.

Remember: The researchers won't be talking directly with any households. Instead, they'll analyse existing information that has already been collected by government agencies. The data will be **de-identified** (names and personal details removed) but **linked** at an individual level, meaning the researcher can see one person's complete journey through their household and area characteristics and income without knowing who they are.

Scenario F: EcoEnergy

EcoEnergy is a small company based in Glasgow that provides advice to homeowners on how to make energy efficiency improvements to their homes. Through their app/website you can enter your postcode, receive a personalised plan, explore grants and funding available, understand future bill savings, compare costs from trusted installers. EcoEnergy works with a range of organisations including the government, lenders, utility providers, employers, local authorities. The service is free to customers, and EcoEnergy makes money through commissions when introducing homeowners to lenders.

EcoEnergy want to identify which properties are least efficient and what the characteristics of these households most commonly look like to target their communication and reach households that would benefit most. They plan to proritise the insulation scheme to households with children, people over the age of 65 and those with specific health conditions and need to plan demand accordingly.

EcoEnergy is requesting access to data about homes in Scotland, including health conditions of residents and age of residents. They already have access to energy efficiency ratings (EPC), property types, property prices, and council tax bandings through open data.

What data does EcoEnergy need?

Detailed property information, including building materials, heating systems, and energy efficiency ratings (linked by Unique Property Reference Number).

Characteristics of area (SIMD) and household (through Unique Property Reference Number) such as age of household members and health conditions

Land ownership, as well as property types and sizes, prices (through the Scottish land registry Local Authority data).

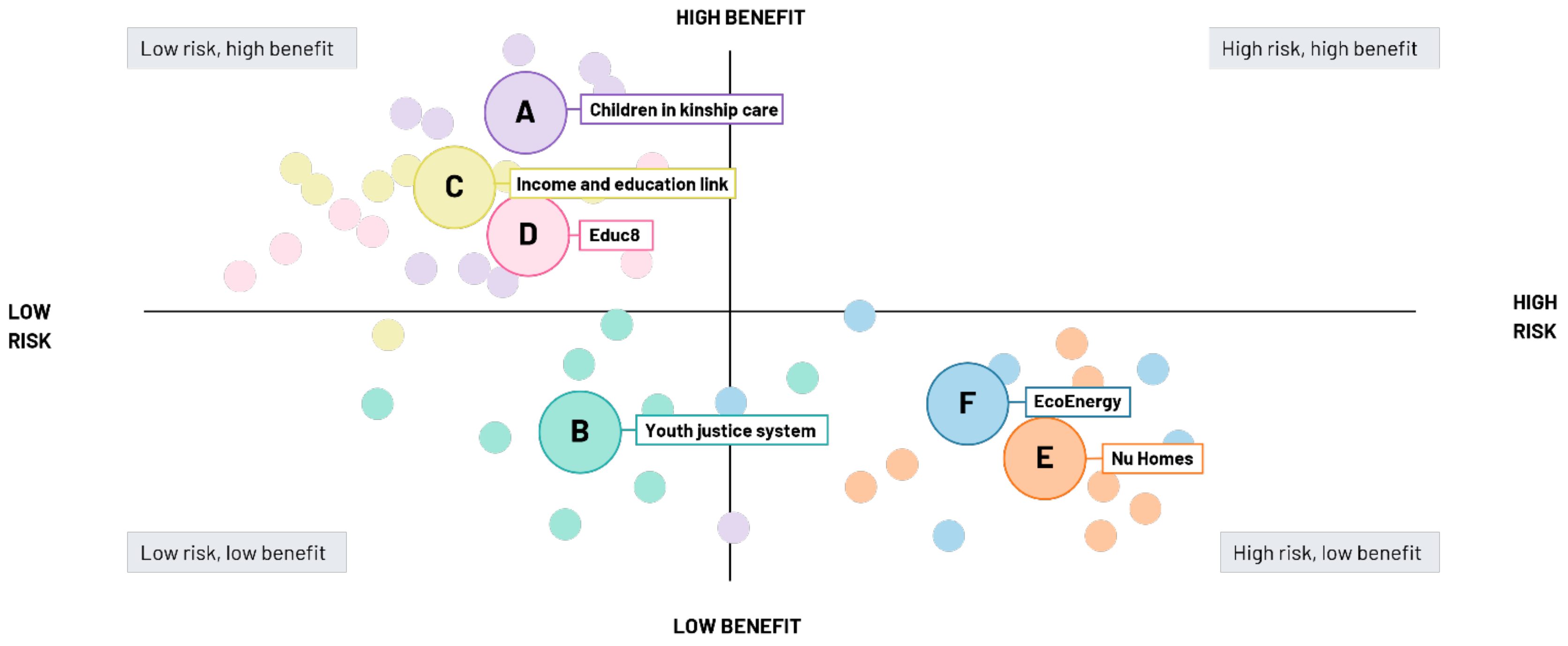
Remember: The researchers won't be talking directly with any households. Instead, they'll analyse existing information that has already been collected by government agencies. The data will be **de-identified** (names and personal details removed) but **linked** at an individual level, meaning the researcher can see one person's complete journey through their household and area characteristics and income without knowing who they are.

F. Risk-benefit quadrant

The risk-benefit quadrant chart below provides an illustration of how participants in the in-person workshop weighed up the risks and benefits of the scenarios they reviewed. Across the six hypothetical scenarios, the children in kinship care, income and education and Educ8 projects were felt to be in the public good, having balanced the relative risks and benefits. The EcoEnergy and Nu Homes projects were not felt to be in the public good, considering the relatively high risk and low benefit. The youth justice system project drew mixed conclusions.

The results presented in the quadrant are based on the post-it exercise conducted with the members of the in-person workshop only. Due to time constraints, the same exercise was not replicated in the online cohort of participants. While not captured in the figure, how participants in the online workshop discussed the risks, benefits and public good in relation to the scenarios are reflected in the main chapter.

Figure 2.8: Risk-benefit quadrant (in-person workshop)

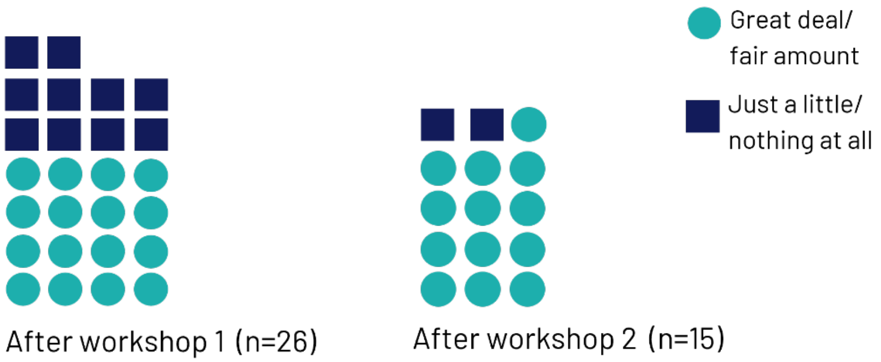


H. Workshop 1 and 2 survey results

After the first and second workshops, participants were sent a short survey to measure understanding of key concepts over the course of the research (see below). The survey results were used to identify any concepts where further clarification was needed at subsequent workshops, but also to gather insights around how messages were landing at key stages. The results suggest that participants' understanding of key concepts improved over the course of the research. However, these should only be read illustratively, as not all participants completed the survey after workshop one, and fewer still completed the survey after workshop two.

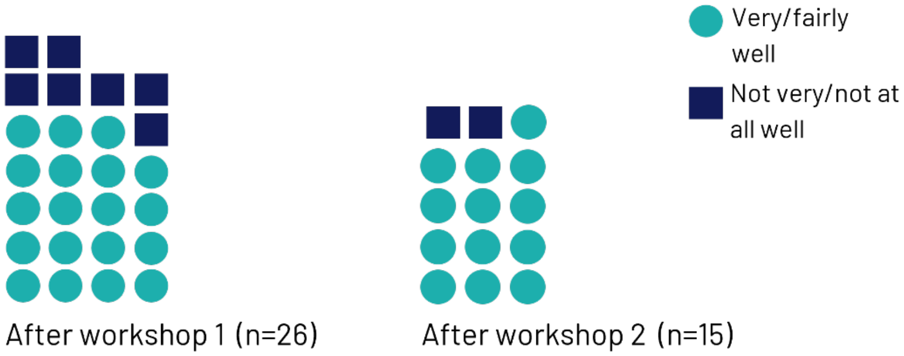
Survey results

How much (if anything) would you say you know about the use of data about people living in Scotland for research?



How well would you say you understand how data is protected?

How well would you say you understand the concept of ‘public good’ when using data about people living in Scotland for research?



How well would you say you understand the purpose of RDS?

Our standards and accreditations

Ipsos' standards and accreditations provide our clients with the peace of mind that they can always depend on us to deliver reliable, sustainable findings. Our focus on quality and continuous improvement means we have embedded a "right first time" approach throughout our organisation.

ISO 20252

This is the international specific standard for market, opinion and social research, including insights and data analytics. Ipsos UK was the first company in the world to gain this accreditation.

Market Research Society (MRS) Company Partnership

By being an MRS Company Partner, Ipsos UK endorse and support the core MRS brand values of professionalism, research excellence and business effectiveness, and commit to comply with the MRS Code of Conduct throughout the organisation & we were the first company to sign our organisation up to the requirements & self-regulation of the MRS Code; more than 350 companies have followed our lead.

ISO 9001

International general company standard with a focus on continual improvement through quality management systems. In 1994 we became one of the early adopters of the ISO 9001 business standard.

ISO 27001

International standard for information security designed to ensure the selection of adequate and proportionate security controls. Ipsos UK was the first research company in the UK to be awarded this in August 2008.

The UK General Data Protection Regulation (UK GDPR) and the UK Data Protection Act 2018 (DPA)

Ipsos UK is required to comply with the UK General Data Protection Regulation (GDPR) and the UK Data Protection Act (DPA). These cover the processing of personal data and the protection of privacy.

HMG Cyber Essentials

Cyber Essentials defines a set of controls which, when properly implemented, provide organisations with basic protection from the most prevalent forms of threat coming from the internet. This is a government-backed, key deliverable of the UK's National Cyber Security Programme. Ipsos UK was assessed and validated for certification in 2016.

Fair Data

Ipsos UK is signed up as a "Fair Data" company by agreeing to adhere to twelve core principles. The principles support and complement other standards such as ISOs, and the requirements of data protection legislation. .

For more information

Ipsos Scotland
4 Wemyss Place
Edinburgh EH3 6DH

Telephone: +44 (0)131 385 1050

www.ipsos.com/en-uk