

# **Public dialogue on the use of data by the public sector in Scotland**

**Findings from a pilot public engagement  
panel for the Scottish Government**

**Report commissioned by the Scottish Government**

**August 2024**

# **Public dialogue on the use of data by the public sector in Scotland**

**Findings from a pilot public engagement  
panel for the Scottish Government**

**Final report prepared for the  
Scottish Government by**

**Ipsos Scotland**



# Contents

<b>Executive summary</b> .....	<b>4</b>
<b>Introduction</b> .....	<b>13</b>
<b>The panel’s starting point</b> .....	<b>18</b>
<b>Looking back: past projects that the Data and Intelligence Network supported</b> .....	<b>23</b>
<b>Looking forward: possible future projects</b> .....	<b>33</b>
<b>Ethical guidelines</b> .....	<b>40</b>
<b>Future engagement</b> .....	<b>47</b>
<b>Conclusion</b> .....	<b>51</b>
<b>Appendices</b> .....	<b>55</b>

## List of figures

<b>Figure 1.1: Guidelines developed by the panel:</b> .....	<b>9</b>
<b>Figure 1.2: Structure of the public dialogue</b> .....	<b>15</b>
<b>Figure 1.3: 3 words to describe the session (online community feedback)</b> .....	<b>49</b>
<b>Figure 1.4: Jamboard exercise for the shielding list project</b> .....	<b>Error! Bookmark not defined.</b>
<b>Figure 1.5: Jamboard exercise for the CURL project</b> .....	<b>Error! Bookmark not defined.</b>
<b>Figure 1.6: Jamboard exercise for the equalities project</b> .....	<b>Error! Bookmark not defined.</b>
<b>Figure 1.7: Jamboard exercise for the Ukrainian Displaced People project.</b>	<b>Error! Bookmark not defined.</b>

## List of tables

<b>Table 1.1: Demographic profile of panel</b> .....	<b>57</b>
<b>Table 1.2: Session summaries</b> .....	<b>60</b>

# Ministerial foreword

With some of the richest data in the world, Scotland is well-placed to unlock the full potential of this asset to drive ethical innovation in our public services and our economy, to achieve better outcomes for the people of Scotland.

Data is the golden thread that enables us to understand and respond to the greatest challenges we face as a society, for instance, driving our cross-sector response to the global pandemic and the Ukraine crisis.

The ethical and innovative use of data supports [our vision](#) to be a wealthier, fairer, greener and more equal country. It underpins our ambition to create an [Ethical Digital Nation](#), where “people can trust public services and businesses to respect privacy and be open and honest in the way data is being used”; where children and vulnerable people are protected from harm; and where data use is safe, transparent and accountable.

Trust in the secure and ethical use of public data is our guiding ethos.

Building on our democratic traditions, we will empower citizens to work with the Scottish Government and our partners across the sector to shape ethical principles that reinforce robust legal safeguards to protect citizens’ data. Their participation, oversight and scrutiny will ensure due diligence, amplify their voice in decision-making and strengthen the legitimacy of digital and data-led activities.

The people of Scotland will be the guardians of our approach to how citizens’ data is used. This is why the Scottish Government and Research Data Scotland convened a pilot public engagement panel to explore the ethical implications of sharing citizens’ data in the public interest.

As you’ll read in this report, the engagement drew out rich and nuanced insights from the public participants on a complex and vitally important subject. There is a strong realisation among the participants of the potential benefits from ethical data-sharing, and the important role the public can play in helping shape how data is used by the public sector.

Reassuringly, the findings highlight a good level of confidence in the safeguarding of data by the Scottish public sector. I also welcome this opportunity to seek public scrutiny on data-led activities that were delivered under the challenging circumstances of recent years, and look forward to learning from this feedback.

I am grateful to all members of the panel who have contributed to this pilot. We will review the ethical guidelines they have produced, and reflect on the emerging findings from this dialogue.

This report also complements our project to [unlock the value of Scotland’s public sector personal data](#), when used with or by the private sector. Taken together, this provides a solid foundation on which to build further engagement with the public to create conditions that enable data to be used in ethical, trustworthy ways for the benefit of all.

**Richard Lochhead, MSP**

Minister for Business

# Preface

The Scottish Government is committed to using data in ethical, transparent and trustworthy ways to deliver better outcomes for the people of Scotland.

The [National Digital Strategy](#) sets out a strategic framework for building an ethical digital nation, where data plays a central role in realising Scotland's ambitions, from addressing climate change to delivering high-quality public services, to achieving sustainable economic growth. These goals align with the Scottish [National Performance Framework](#) outcomes; specifically Human Rights, Children and Young People, International and Communities.

The Scottish Government's vision is of a society where people are empowered to control their personal information and actively participate in the design of products and services that better meet their needs.

Inspiring public trust in data use requires public participation in decision-making. Public bodies will secure and maintain a social licence to operate, by embedding the voices of citizens, and reflecting the diversity of public views, in their data projects.

Mobilising data has driven the Scottish Government's rapid, cross-sector response to major challenges such as Covid-19 and the Ukraine crisis, and catalysed innovative approaches to improving people's health and societal wellbeing.

Post-pandemic, the Scottish Government is focused on leveraging ethical data use to address a broader range of challenges. Our aim is to build on effective approaches, and to work with the public and cross-sector partners, to shape ethical guidelines that the Scottish public sector can operationalise and, in so doing, unlock the value of this data, for societal benefit.

Foundational to this work is understanding what the ethical use of data means to Scottish people in a fast-evolving, increasingly complex, post-pandemic world. Together with Research Data Scotland, the Scottish Government convened a public panel to explore the ethics of a range of high-profile data projects.

The purpose of the panel was to inform approaches to data use by the Scottish Government and public sector bodies in Scotland. The public dialogue that underpins the findings presented below was sponsored and co-managed with [Sciencewise](#) – a public engagement programme led by the UK Research and Innovation which enables policy makers to develop socially informed policy, with a particular emphasis on science and technology.

The findings presented in this report will inform policies in support of the wider public sector commitment to ethical and transparent use of data about citizens.

# Executive summary

## Background

The Scottish Government has stated a commitment to using data and digital technologies in an ethical way for the benefit of the people of Scotland.<sup>1</sup>

During the pandemic, the Data and Intelligence Network (DIN<sup>2</sup>) was set up by the Scottish Government as a dedicated team providing additional capabilities to ensure that data were utilised effectively and ethically to address key challenges relating to COVID-19. The need for urgent and decisive action during the COVID-19 pandemic meant that data-led projects were fast-tracked to implementation, and opportunities for consideration of public views on the social and ethical implications of the data used were limited.

As Scotland moved out of the pandemic, the DIN's resources were redirected within the wider Scottish Government to address cross-cutting data challenges and to advise on ethical dilemmas relating to the use of data, and to support the recovery process from the pandemic.

The Scottish Government together with partners from Research Data Scotland and UK Research and Innovation's Sciencewise programme, decided to convene a public panel to explore the ethics of different data-led projects, including those that took place during the pandemic. The purpose of the panel was to inform approaches for data use by the Scottish Government and public sector agencies in Scotland.

## Approach

Ipsos and its partners at the University of Edinburgh designed and facilitated a public panel that brought a group of 25 people from across Scotland together to explore perceptions and understanding of public sector data-led projects. At the end of the dialogue, the panel produced a set of ethical guidelines to inform the way the Scottish Government and public sector organisations use data about citizens. A wider aim of the project was to reflect on this process and provide lessons for possible future engagements involving the public in data policy, scrutiny and decisions.

Using a deliberative public dialogue approach, the panel met over six three-hour online workshops between September and December 2022 to answer the question: ***What guidelines should the public sector follow when using citizens' data?*** The panel listened to presentations from experts, learned about the issues, and discussed them together before drawing conclusions to form a set of ethical guidelines. In March 2023, 15 members of the panel were reconvened for two additional workshops to further explore issues that were not covered in

---

<sup>1</sup> Link to Scottish Government report: [A changing nation: how Scotland will thrive in a digital world \(2021\)](#)

<sup>2</sup> The Data Intelligence Network was the collective name for a network of around 300 organisations, the majority of whom were in the public sector. The DIN team involved a core group of people who delivered services for the Network and Scottish Government employees.

detail in the first six workshops (related to data sharing outside of the public sector). The findings from these additional workshops are provided in a separate report.

## Key findings

### Trust



**The panel were generally trusting of the public sector's use of data.** This trust was linked to an understanding that use of data by the public sector can have benefits for society. It was based on an expectation that the public sector would follow rules and regulations around the use of data and be held accountable for any misuse. Participants wanted reassurance that the public sector would follow ethical principles alongside existing legal frameworks.

### Clarity of purpose



**Participants felt that use of data by the public sector should have a clearly defined purpose and scope,** to avoid misuse of data about citizens. Participants wanted to see transparency and openness on the reasons for public sector use of data, how that data would be used, and who it would be used by. Underlying this need for reassurance was a broader concern about privacy, and a desire to avoid personal information about individuals being accessed unnecessarily.

### Public benefit



**The panel agreed that the use of data was only acceptable if there was a clear public benefit, or public good.** Public benefit was considered a subjective concept, but was described by participants as “something that benefits society”, “something that improves the lives of individuals” and contributes to “a happy society”. It was felt that public benefit could apply to the whole of society or to a small part of the population (e.g. a minority group).

### Data quality



**Data quality was seen as an important ethical consideration.** Participants stressed the importance of data-driven decisions being made with up-to-date and accurate information. Data quality was linked to fairness, with the panel feeling that gaps in data could lead to individuals being excluded or not benefitting from certain initiatives. They therefore wanted to see a minimum quality standard put in place which future data-led projects would be required to meet.

### Urgency



**Views were influenced by the context in which data were being used, specifically whether it was an emergency situation or not.** This public dialogue gave participants the opportunity to scrutinise and share feedback on real public sector data-led projects, some of which had been delivered under the unprecedented circumstances of the COVID-19 pandemic. Participants understood that data had to be used quickly to support decisions related to COVID-19 and they recognised the benefits of doing so. They therefore identified a need for some flexibility to be allowed for in the ethical guidelines, while adhering to basic principles.

**Avoidance of harm**



**The panel was sensitive to the impacts of data use, both positive and negative, on marginalised groups.** Protection of individual privacy was seen as particularly important when dealing with sensitive, special category data and the panel had concerns that misuse of this type of data could lead to individuals being discriminated against.

**Accountability**



**Participants felt that an independent body should oversee decisions about data use** and hold organisations accountable for any misuse of data. It was suggested that this could be in the form of an independent panel.

**Involve the public**



**The panel felt that the public had an important role to play in helping shape how data were used by the public sector.** There was overwhelming support for future public engagement on the use of data and participants felt that a public panel, would be a good way of engaging the public on this topic in future.

### Ethical guidelines for public sector use of citizens' data

The guidelines developed by the panel are outlined overleaf, grouped under six key themes (purpose, transparency, public benefit, accountability, data quality and context).

#### Figure 1.1: Guidelines developed by the panel:

**When using citizen's data, the public sector should manage the PURPOSE by:**

- Ensuring the purpose for using the data is clearly defined and data is used only for that purpose. Timescales for use should be clearly defined.
- Having a clearly agreed justification for using citizens' data (i.e. if there is a clear public benefit) and ensuring that only data that is necessary for the project is used.
- Ensuring that data are not used solely<sup>3</sup> (directly or indirectly) for profit by private sector organisations. The public sector should ensure that it and private sector partners only use data proportionate to the specific purpose it was collected for.
- Not using data outside the scope of any consent that applies to the data.
- Not sharing data beyond the agreed organisations. If more organisations are included later in a project, they should go through an ethical assessment.

---

<sup>3</sup> Inclusion of the word 'solely' was not agreed upon by all participants in the final ratification of these guidelines, but rather reflects general discussions around the involvement of the private sector in using data about citizens. This explained in more detail in the "deliberative journey" chapter.

**When using citizen's data, the public sector should ensure TRANSPARENCY by:**

- Making clear what data are being used and for what purpose.
- Making clear which organisations can access the data, and why.
- Specifying how long data will be stored for before deletion.
- Ensuring the public can easily access information about the project, including: what data are being used and for what purpose, how long data are stored before they are deleted, and a summary of findings or impact of project (where it is legally possible to do so and where individuals are not identified).

**The public sector should ensure the use of citizen's data is in the PUBLIC BENEFIT by:**

- Clearly defining and explaining what the public benefit is.
- Considering whether the public benefits of using the data clearly outweigh the risks. Any potential harms from use of the data need to be analysed and weighed against the benefits.
- Considering negative impacts to the public and/or the environment or economy, with possible longer term impacts also considered. Projects that benefit or make a positive impact on a small number of people can be in the public benefit, provided they do not negatively impact others, the environmental or the economy.
- Ensuring that identifiable data are only used if it meets the standard of achieving public benefit.

**When using citizen's data, the public sector should ensure there is ACCOUNTABILITY by:**

- Clearly documenting the process used to decide whether the project should go ahead (to an agreed formal structure)
- Ensuring there is a hierarchical organisation chart to show who is responsible/accountable for each aspect/stage of the project.
- Seeking approval and oversight from an independent panel on whether a data project should go ahead or not, including whether public benefits outweigh risks. The panel should make decisions based on what is in the best interests of the public and there should be no declared conflicts of interest on the panel.
- Consulting members of the public on the acceptability of the use of the data (for determining principles but not to decide if a project should go ahead or not – this is the role of the independent panel).

- Ensuring an ethical assessment is carried out once the scope of the project is known.
- Taking responsibility when something goes wrong and stopping the project if necessary.
- Ensuring there is independent oversight from a third party (e.g. Information Commissioners Office and DIN) for projects involving the private sector, with clear sanctions for misuse (criminal and civil).

#### **When using citizen's data, the public sector should ensure DATA QUALITY by:**

- Establishing and publishing a minimum quality standard for data projects (that includes consideration of how much data is needed). The extent to which data projects meet the threshold for data quality must be checked and continually assessed by the team delivering the project. If there is involvement from the private sector, these checks should be made by someone from government/public sector.
- Using up to date data that matches the agreed purpose and specific scope.
- Ensuring data are held securely for an agreed period after a project to allow for quality checking.
- Determining who can access the data and monitoring who has accessed the data.

#### **When using citizens' data, the URGENCY should be considered, by:**

- Defining what constitutes an emergency. Any impacts of flexing guidelines in this context should be assessed continually, as far as practical, and after the fact (including any lessons learned).
- In an emergency situation, such as where there is threat to life, it may be necessary for data to be used that was not part of the original scope. Considering whether the public benefits of using the data clearly outweigh the risks.
- In the event of an emergency the use of identifiable data can be justified. If the private sector is involved, there should be clear rules about what private sector organisations do with data after an emergency including when they are deleted.
- In an emergency situation, it may be necessary for the timescales for data retention and deletion to be reviewed and extended.

#### **Future engagement**

Participants felt that the public had an important role to play in helping shape how data was used by the public sector and supported a panel-style approach.

One of the aims of this public dialogue was “to create a blueprint for a long-term, sustainable form for engaging and involving the public in data policy, scrutiny and decisions”. There was overwhelming support for future public engagement on the use of data, which reflected participants’ positive feelings about their own experiences of this process.

A unique aspect of this public dialogue was that participants acted as a panel, meaning they had the opportunity to review and appraise past, current and future data-led projects. The panel-style approach placed participants in the role of evaluator, giving feedback (sometimes directly to those involved) on data-led projects in a way that could influence decisions around their future delivery. The panel-style approach (with the group meeting over a four-month period, and gaps of up to two weeks between workshops) also meant that they had a fairly long period of time to immerse themselves in the topic, reflect in between sessions and gradually develop their ethical guidelines in response to what they learned.

Participants felt that a public panel, designed and structured in a similar way to the one they were part of, would be a good way of engaging the public on this topic. Potential uses for a panel could be to review and provide feedback on potential data-led projects, or to revisit ethical guidelines developed by this panel to test whether they were still appropriate. Other suggested forms of engagement included teaching children about data at school, and using websites (such as public sector websites) to invite feedback from the public about potential data-led projects.

Overall, it was felt that members of the public could provide a balance to the views of subject matter experts and data specialists, potentially offering new ideas or alternative issues to inform future data use and wider policy. If the data in question had originated from members of the public, it was seen as only fair and transparent for the public to have a say in how those data would ultimately be used.

# Introduction

## Background

The Scottish Government has stated a commitment to using data and digital technologies in an ethical way for the benefit of the people of Scotland.<sup>5</sup>

During the pandemic the Data and Intelligence Network (DIN) was set up by the Scottish Government as a dedicated team, providing additional capabilities to ensure that data were utilised effectively and ethically to address key challenges relating to COVID-19. The DIN operated from within the Scottish Government, providing skills and expertise to a wide range of organisations across Scotland's public sector and providing support with their data-led projects at the height of the COVID-19 pandemic. The projects that the DIN advised on presented ethical dilemmas in how to reconcile the privacy, rights and freedoms of people in Scotland with rapid, proactive, and responsible use of information. The need for urgent and decisive action during the COVID-19 pandemic meant that data-led projects were fast-tracked to implementation, and the opportunities for consideration of public views on the social and ethical implications of the data use were limited.

In an increasingly data driven world, good governance is required to ensure public data are used effectively, ethically and appropriately. Engaging the public – those whose data is used in research, planning and service development and delivery – helps to develop trustworthy and robust frameworks for how government and other agencies collect, analyse and use data. Public engagement can take many forms and have multiple purposes, but an overriding aim is to promote better policy and decision making, often through deliberative engagement.<sup>6</sup>

The Scottish Government, together with Research Data Scotland (RDS), decided to build on the work done by the DIN which had started to expose some of the ethical challenges of data-led projects. The Scottish Government and RDS agreed to convene a public panel to explore the ethics of past data-led projects supported by the DIN, and possible future projects to inform approaches to data use by the Scottish Government and public sector agencies in Scotland.

## Research objectives

The aim of this public panel was to explore perceptions and understanding of public sector data-led projects in order to produce a set of ethical guidelines that the Scottish Government and public sector organisations should follow when using data about citizens. The key research objectives of the panel itself were to explore:

---

<sup>5</sup> Link to Scottish Government report: [A changing nation: how Scotland will thrive in a digital world \(2021\)](#)

<sup>6</sup> Aitken, Mhairi, Mary P. Tully, Carol Porteous, Simon Denegri, Sarah Cunningham-Burley, Natalie Banner, Corri Black, et al. 2019. 'Consensus Statement on Public Involvement and Engagement with Data-Intensive Health Research'. *International Journal of Population Data Science* 4 (1). <https://doi.org/10.23889/ijpds.v4i1.586>

- Public perception and understanding of public sector data-led projects (involving different types of data including anonymised or identifiable records from NHS health data, census data, education data, housing data, and location data).
- Levels of public trust in different methodologies to ensure privacy of individuals' data.

Broader objectives of the project were to:

- Drive an increase in the amount of public engagement on data use taking place across the public sector by introducing and involving a range members from the DIN Network in the design and delivery of the project.<sup>7</sup>
- Create a blueprint for a long-term, sustainable forum for engaging and involving the public in data policy, scrutiny and decisions.

The public dialogue reported here builds on previous public engagements in Scotland on the use of data (key insights from which are summarised in appendix A).<sup>8</sup>

## Methodology

Ipsos, along with its partners at the University of Edinburgh, designed and facilitated a public panel that was funded and guided by the Scottish Government and UK Research and Innovation's [Sciencewise programme](#), an internationally recognised public engagement programme which enables policy makers to develop socially informed policy. The specific methodology used with the public panel is known as a "public dialogue".<sup>9</sup> Public dialogue is a process during which members of the public interact with scientists, stakeholders and policy makers to deliberate on issues relevant to future policy and research decisions.

The panel brought together a group of 25 people from across Scotland to learn about the topic of data use by the Scottish Government and public sector agencies. The panel met over six three-hour online workshops<sup>10</sup> between September and December 2022 to answer the following key question:

***What guidelines should the public sector follow when using citizens' data?***

---

<sup>7</sup> This objective was set at the beginning of the process, but the nature of the Network membership (consisting of over 300 organisations) meant that it was not practical to engage widely across the Network on the design of this public dialogue. However, Network members were involved in the project team and the Oversight Group, contributing to review of discussion guides and stimulus materials and delivered presentations to the panel.

<sup>8</sup> A rapid evidence review was conducted to inform the design of this public dialogue. This review synthesised the published results of public engagement work from 2011 onwards that focussed on the use of data.

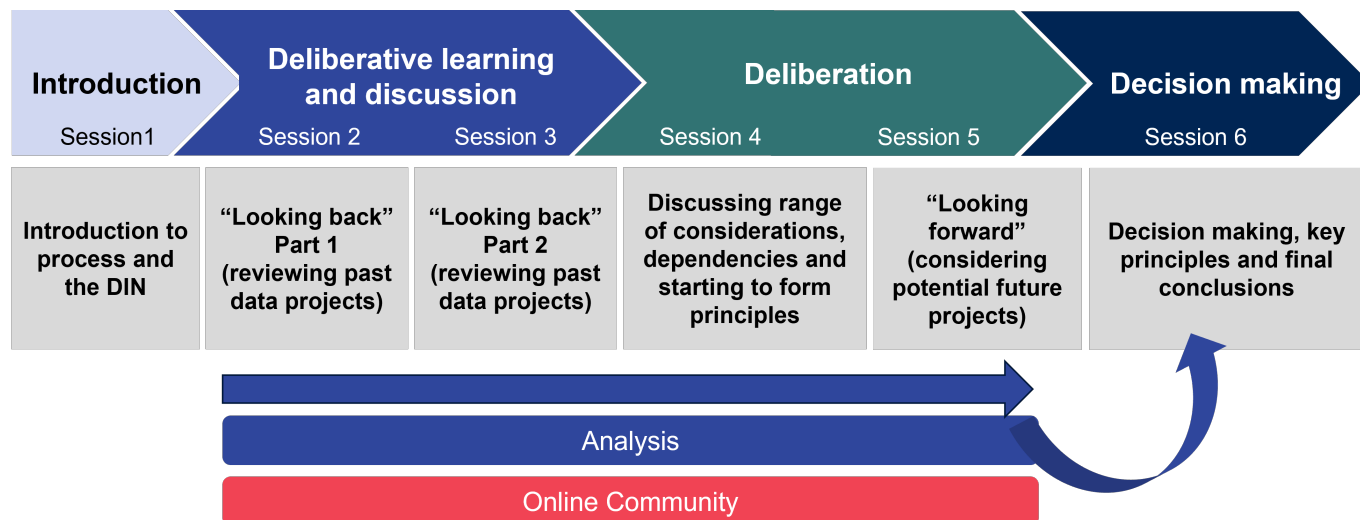
<sup>9</sup> <https://sciencewise.org.uk/about-sciencewise/our-guiding-principles/>

<sup>10</sup> A smaller group of participants met for an additional two workshops to explore the use of data by private and third sector organisations and the idea of benefit sharing (the findings of which can be found in a separate report).

Over the course of the public dialogue, participants reviewed different data-led projects that the DIN had been involved in previously or were considering involvement in. The panel listened to presentations from experts, learned about the issues, discussed them, and then drew conclusions together to form a set of ethical guidelines which are presented in this report.

Further details about the process (including an overview of each session with dates, times, content and specialists) can be found in appendix D but the overarching design of the dialogue is summarised in figure 1.2 below.

**Figure 1.2: Structure of the public dialogue**



Alongside the online meetings, an online community helped support ongoing engagement with panellists, facilitating continued discussion and reflection.

### Oversight

An oversight group – comprising Scottish Government representatives, Research Data Scotland representatives, and independent experts from academia and civil society – met regularly to advise on the methodology. The oversight group provided checks and challenges over the course of the project to ensure the design of the panel was appropriate, relevant and robust. A list of oversight group members and meeting times is detailed in appendix B.

### Sampling and recruitment

Participants were recruited using a civic lottery approach,<sup>11</sup> which involves inviting a random selection of households to participate. They were recruited to be broadly reflective of the Scottish population in terms of age, gender, region, ethnicity, disability and education. Ethnic minority groups were over-sampled to ensure sufficient representation of these groups. An attitudinal measure was also included in the selection process to ensure a range of views were

<sup>11</sup> Link to the Sortition Foundation website: [How to run a citizen's assembly](#)

represented in terms of trust in the Scottish Government and public sector agencies to use data for the public good. The recruitment process is outlined in more detail in appendix C.

Overall, 30 people were selected to join the panel and 25 participated throughout. A table summarising the demographic profile of the final selected and confirmed sample can be found in appendix C.

To support and enable participation in all workshops, participants were each paid £400 for joining the online sessions and online community. Where necessary, participants were provided with equipment, such as headsets, laptops or internet dongles and were supported with training on how to use the technology and access the meeting platform.

### **Materials and input from specialists**

Discussion guides and stimulus materials were developed by Ipsos and reviewed by the Scottish Government, Sciencewise and the oversight group. A range of specialists joined at different points in the dialogue to provide information that would be useful for participants' learning and deliberation. In the first session, specialists included academics and professionals who explained key concepts to support participants' discussions (including the role of the DIN, the legal context of data use and an introduction to data ethics). In the following sessions, those involved in the data-led projects being reviewed by the panel provided overviews of these, while academics and representatives from independent research institutes provided alternative perspectives on the projects to help participants consider different views on the use of data about citizens.

Presentations were either delivered live or recorded in advance and played live during the sessions. Some specialists presented in plenary and then stayed to answer questions that arose during breakout discussions. Others presented in smaller groups, remaining in the groups afterwards to take questions directly from participants. Any questions that were not answered during the live sessions were compiled in a Question and Answer (Q&A) document (see appendix E).

### **Interpretation of qualitative data**

The guidelines set out and discussed in this report are intended for consideration in the possible future use of data about citizens by the public sector.

This report synthesises the diverse expressions of participants to draw out major themes of discussions and to draw attention to the way that the panel – individually and collectively – made sense of a complex topic, describing what mattered to them and why. On occasion, 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 important about how the issues can be perceived and understood by members of the public.

A robust and systematic analysis approach was used, with conclusions based on groups that are reflective of the diversity of the wider public. The deliberative nature of the project allowed for ongoing analysis throughout fieldwork, which ensured that emerging principles and themes -

both from workshop discussions and online community activities - could be played back to participants as the dialogue progressed. Analysis does not seek to quantify findings nor does it indicate statistical significance from a representative sample. A more detailed summary of the analytical approach to the dialogue can be found in appendix D.

This report offers a constructive insight into public perspectives on the key questions posed to them after receiving and deliberating on essential information relevant to the questions.

## The panel's starting point

This chapter outlines the various entry points at which participants joined the panel, their reactions to the initial information provided, and the sorts of questions they raised. It provides the baseline against which later findings can be compared, as participants' views on key topics developed over the course of the dialogue. It also highlights useful lessons for future engagement on public sector use of data.

Members of the panel were recruited from across Scotland and it was emphasised that they did not need any prior knowledge to take part, just a willingness to listen and share views. Participants therefore came with varied knowledge, interests and understanding in relation to the use of data about citizens by the public sector.

In session one, the panel started learning about some of the key concepts that would help them in later deliberations. They heard three presentations which introduced the DIN, outlined data protection and the legal context surrounding data use, and summarised some key concepts in relation to data ethics.

### Varied expectations and starting points

At the start of session one, participants were asked to share their hopes and fears about the process that lay ahead. Expectations varied, but common words used were “curious”, “interested” and “intrigued”. Topics they were interested in finding out more about included: how data were protected, how the public sector uses data and why, and the extent to which “personal” or “private” information was accessed by the public sector.

Participants had different starting points in terms of their experience of and interest in data. Some were involved in aspects of data protection in their jobs, for example through working in IT or being data controllers for their projects or clients. Some framed their experience of data in terms of their online lives, for example mentioning that they were concerned about their online footprints and that they disabled cookies or targeted ads to reduce the amount of data being collected about them. Others brought no prior experience or shared any specific perspectives on data. Across the board, however, there was an appetite to learn and explore the topic further.

“I was interested in taking part in the project as data is very important. It's important to understand how data is used and what parts are kept and passed on to other organisations...If organisations have too much data, that can be a concern.” (Session one)

Participants also shared expectations about their involvement in the panel itself and, while there was some trepidation about what it would involve, they generally felt positive about the opportunity to share their views and have their voices heard.

## Broad trust in the public sector from the outset

Participants started the dialogue from a broad position of trust in the public sector using data about citizens, but with mixed levels of awareness of how data were used and the legal frameworks that underpin data use.

Before they heard from any external specialists, participants were asked the extent to which they trusted the Scottish Government and the public sector to use data about citizens for the public good. The dominant sentiment at this early stage was one of trust. This was based on the assumption that strict guidelines were in place to govern public sector use of data. It was also driven by a broader sense that the Scottish Government exists to deliver services that are for the public good.

“While data can be misused, you have to assume the majority of people in government are going to want to try to move to something better, helpful and easier for people, and to generally do good things rather than bad things”. (Session one)

Among the few participants who were less trusting in public sector use of data, reasons for this related to a general distrust in government and a feeling that data were “over-shared” and that organisations (not just the public sector) have too much access to individuals’ data. Concerns about potential data breaches were raised, including examples from outside the public sector, such as the Facebook-Cambridge Analytica scandal where the personal data of millions of Facebook users was collected without consent.<sup>12</sup> There was also a sense that the public are not well informed about when and how their data were used, which leads to mistrust.

“I don't think it's that transparent about what data they are collecting and what the purpose is, why they're collecting it. And along with that, how secure and how do they manage to make sure that it's always up to date? I guess I don't have visibility of any of that”. (Session one)

Trust in public sector use of data was framed relative to that of the private sector, of which participants were more cautious. Early in the dialogue, there was less confidence in the safeguarding of data by the private sector than the public sector. There was a perception that government and the public sector took data security seriously, but that standards were not as high or as consistent within the private sector (the findings of the additional workshops which explored private sector involvement in more detail can be found in a separate report).

“From the public sector, I feel fairly confident that there is safeguarding in place...I'm interested in different organisations that may not follow the same safety standards.... Maybe [private sector] organisations need to go through ethical standards training on handling public information.” (Session one)

---

<sup>12</sup> Link to BBC News webpage: [Facebook-Cambridge Analytica Scandal](#)

## Mixed awareness and understanding about the legal context

Participants came with mixed levels of awareness of how data was used and the legal frameworks that underpin data use. In session one, participants were given an overview of the UK General Data Protection Regulation (GDPR) and the Data Protection Act (2018) and the key principles of the legislation. Prior to hearing these presentations, participants' awareness of existing data protection legislation had varied, from those who were knowledgeable about data protection law through their own jobs, to those who said these were relatively new terms for them.

Participants considered the legal context to be complex and described the information presented to them as dense and difficult to digest. However, the existence of the legislation provided a sense of reassurance that a system was in place to protect data. Navigating the relationship between the legal and ethical issues associated with data use was an ongoing challenge for the panel as deliberations progressed.

One example of this was the issue of consent. Consent was introduced to participants in session one, both from a legal perspective as a lawful basis for processing personal data under GDPR and from an ethical perspective in terms of balancing individual and public interests. Participants revisited consent at different points throughout the dialogue and there was an ongoing struggle to reconcile consent (i.e. the importance of individuals having a choice about their data being used and the right to privacy) with other issues, such as the quality of data (i.e. gaps in information leading to poor policy decisions) or the context (i.e. needing to respond quickly in an emergency).

## Low awareness of levels of identifiability in data

In the first session, participants learned about the different ways in which data was collected, the form it could take, and the varying levels of identifiability in data. There was low awareness initially about what was meant by the different types of data – anonymised, pseudonymised, or de-identified – and the extent to which individuals would be identifiable. Participants asked for clarity on these concepts, and over the course of the dialogue participants continued to revisit these terms and clarify their understanding.

### A question raised by the panel in session 1:

- “When data is shared for research purposes, is it de-identified? And have people consented to their data being shared for this purpose?”

## Lack of clarity around the role of the Data and Intelligence Network

The information provided in the presentation about the DIN gave some reassurance that there was a system in place to oversee public sector use of data, in particular the ethical framework<sup>13</sup> that the DIN expected members of the Network to adhere to when running their data projects.

However, following the presentation, participants were still not clear about the role of the DIN. The panel asked a range of questions to build their understanding about the DIN's purpose, structure, and the extent of its influence:

### Some questions raised by the panel in session 1:

- “Where does the DIN fit within the public sector – are they part of the Scottish Government? Are any private sector organisations involved in their work?”
- “Is the network mainly just in place to manage ethical standards, or are they a group collective that decides what [data] should be shared between organisations?”
- “How are decisions made about what kind of data is shared with which kind of organisations?”

Uncertainty about what the exact role of the DIN was persisted for some participants throughout the dialogue, as the distinction of roles and responsibilities between the DIN team and other members of a data-led project team were not clear to them.

### Questions over the role of the panel

The presentation on data protection generated further questions around how decisions were made in relation to data ethics and who made those decisions. This also led some to question the role of the panel:

### Some questions raised by the panel in session 1:

- “If data sharing should be done ‘to serve mankind’, who decides what mankind is?”
- “If GDPR and DPA set policy, what is our role in this public panel?”

The presentation on the ethical issues helped to settle this in participants' minds to some extent, with the understanding that something might be legally acceptable but not morally acceptable. A further recap on the role of the panel was presented at the beginning of session two; feedback suggested that this made the purpose clearer and enabled the panel to move forward with their deliberations, which included thinking about *who* should make decisions about data use.

---

<sup>13</sup> Link to the Data and Intelligence Network Ethics Framework: [An Ethics Framework for the Data and Intelligence Network](#)

The panel's appreciation for the complexity and subjective nature of data ethics evolved as they became more familiar with different types of projects.

The following chapters summarise the key findings from the panel's review of data-led projects that the DIN had been involved in previously, or were considering their future involvement in.

# Looking back: past projects that the Data and Intelligence Network supported

This chapter summarises reflections on four past projects that were presented to the panel. By exploring reactions to and perceptions of specific data projects, this chapter highlights the ethical considerations around data use that were important to participants, which later fed into the guidelines they developed.

These sessions represented a distinct stage in the learning process. By hearing from representatives of real-life data projects, participants had the chance to place some of the concepts outlined in session one (data sharing, data protection, data ethics) into a practical context. It also gave participants the chance to scrutinise projects and ask questions directly to the specialists in plenary Q&A sessions.

The past data projects reviewed in sessions two and three were:

- Shielding list (session two) – Medical records were used by NHS Scotland during the pandemic to identify those more likely to be clinically at risk from COVID-19.
- CURL (session two) – Health data was linked with residential addresses to improve understanding of health risks in different situations.
- Equalities (session three) – Information from medical records, education records and census data was used to develop as complete a picture as possible of the protected characteristics across Scotland.
- Ukrainian Displaced People (session three) – Data was processed and shared during the Ukraine Crisis so that Ukrainians could be safely housed across Scotland.

As well as representatives from each project sharing their reflections, in session three an academic from Tilburg University, Dr Anuj Puri, joined to offer his reflections on the projects from an alternative perspective (having not been involved in them himself). This perspective provided an opportunity for participants to consider other points of view on the ethical issues around the use of data by the public sector.

In this chapter, each project is presented separately, summarising the key ethical considerations raised by the panel as part of their assessments. These reflections formed the basis of the ethical guidelines that the panel developed in later sessions.

## Key findings

- On the **shielding list** project, the panel felt the benefits of saving lives during the pandemic offset the risks and challenges (recognising the potential harms around asking people to shield and adding them to the shielding list without their consent).
- On the **CURL** project, there were more mixed views on the benefits of the project – some felt the linked data could be useful in future while others highlighted a lack of transparency around these possible uses.
- On the **equalities** project, the benefits of having this data available for future use were broadly recognised. The challenges associated with it largely hinged on data quality concerns and the risk of this leading to poor policy decisions based on skewed data.
- On the **Ukrainian Displaced People** project, the humanitarian aspect was broadly applauded. However, there were also concerns raised about data sharing between countries in the context of war.

The **key ethical considerations** raised in relation to these past projects included:

- Ensuring accurate and up to date data.
- Proportionate use and not going beyond the original scope.
- Weighing up the relative benefits and harms to society.
- Ensuring transparency and accountability in decisions about what data is used, by whom, and for how long it is held.
- Ensuring data is held securely.
- Ensuring the principles of consent are adhered to.

## Past project one: Shielding list

During the COVID-19 pandemic, medical records were used by NHS Scotland to identify citizens who were more likely to be clinically at risk from COVID-19. This data were then used to contact individuals to request that they stay in their homes and take extra precautions to minimise their risk of contracting Covid. The data were shared with local authorities so that they could provide additional support to any individuals who were shielding. A summary of this project was presented in plenary by the DIN team, followed by smaller breakout discussion, and Q&A with the DIN team in plenary.

### Strong positive impact of the shielding programme

The shielding programme was seen as having positive impacts by helping keep people safe during the pandemic. Participants reflected on how the shielding list had impacted on their own lives or those they cared about and largely felt the risks associated with sharing identifiable health data were outweighed by the benefits of protecting vulnerable groups.

"I have four friends, all of whom were shielding. I think the government got it spot on, and quickly." (Session two)

### Potential risk of harm

The panel were mindful of the potential harms associated with receiving a letter and being advised to stay at home (such as negative impacts on peoples' mental health and wellbeing). They also highlighted the potential risk to individuals' privacy. For example, one participant expressed discomfort about being on the shielding list and having such information about them shared. This concern related to the risk, identified earlier by the panel, that data could be used for purposes other than shielding.

The project also sparked some discussion about the issue of consent. One view was that information on peoples' health conditions should not be shared beyond the NHS – such as with charities – without their permission. Another view was that it was acceptable to share this information with organisations who could provide support to those that were shielding, as the data may not need to include detailed information (i.e. their name and address but not their health condition). A more exceptional view was that individuals should have been consulted about being included on the shielding list in the first place.

"[A] negative would be maybe the possibility of intrusion, if that is the right word. You don't want someone to know something and you get a letter discussing that. It could be an issue for someone they have to personally deal with." (Session two)

### Importance of data being used for a specific purpose only

Questions were raised about how long the shielding list data would be held for. In discussions around this, participants highlighted the importance of the data not being kept longer than was needed and only being used for the specific purpose of shielding. On balance, the panel was reassured that data had been used proportionately and appropriately, and that there was a clear justification for its use in this specific case.

"I'd be very nervous about day-to-day sharing of data unless it was for a really important purpose, like shielding." (Session two)

### Risk of gaps in the data

Concerns were also raised about data quality, including how accurate and up to date the data were. They noted that gaps in the data may have led to people being missed from the shielding list.

"That's the issue about the data being in the right place and up to date. It's fine if you are keeping your data up to date in the right place, but how do you know if you have missed someone out?" (Session two)

### Necessity of clear roles and responsibilities when multiple organisations are sharing data

Participants raised questions about the number of organisations involved in the shielding list project and were unclear about who was accountable:

### Questions raised in relation to the shielding list project:

- “Who is making the judgment calls about who data is transferred to?”
- “Why are the third party organisations being involved in data sharing?”

Given the range of public sector organisations involved in the shielding list data project (such as health boards, GPs, universities and local authorities) and the range of data sources (such as GP, local authority and academic datasets), the panel felt that clarity over roles and responsibilities was important.

“Most of all, I think that it should be very clear what is being taken and who is getting this data, who it's being shared with.” (Session two)

Participants also felt that the public benefits of projects like the shielding list should be clearly defined and communicated by the organisations involved (not through “long T&Cs”) along with assurances that data were being used responsibly. Having such transparency was linked to building trust in public sector use of data about citizens.

“There's no feedback or follow-up on how it's been used and its impact. If data was used for the good of society, and we knew that, we might trust the organisations more with the data.” (Session two)

In concluding discussions on the shielding list project, the panel noted down their key ethical considerations on a digital whiteboard using post-it notes. .

### Past project two: CURL

Public Health Scotland and academic researchers from the Scottish Centre for Administrative Data Research (SCADR) undertook work to link health data (using Community Health Index – or CHI – numbers) and residential addresses (using Unique Property Reference Numbers or UPRNs). The project was called CHI/UPRN Residential Linkage (CURL). During the pandemic, this project helped the Scottish Government understand the impact of hospital discharges to care homes in terms of COVID-19 outbreaks and improve testing in care homes. A wider aim was to combine this linked dataset with other data for future uses, for example combining it with geography or area-based datasets to understand the impact of flooding on peoples’ health. A summary of this project was presented to the panel in plenary by a DIN team member, and was followed by smaller breakout discussion and Q&A with the DIN team member in plenary.

### Concerns about widening the scope of the project in future

The primary purpose of this data-led project, to minimise the spread of COVID-19 in care homes, was recognised as a positive one. Participants felt that the “tidying up” of data for future use beyond the pandemic would also be beneficial, for example by helping to understand public health needs at a local level. However, some participants were not clear on the possible benefits of linking such data and what difference it could make in the future.

“The care home scenario was a great use of it, but he was talking about bringing this forward into the future... I don't know how things like insulating the roof, like he said, can have an impact on your overall health.” (Session two)

The scope of the project was therefore viewed as a challenge, given the range of possible future uses that were outlined in the presentation. These possible uses – such as for understanding the impact of flooding on peoples' health – were not widely recognised as being relevant to people's health data and were described as potentially “intrusive”. Although it was deemed appropriate to link this data to protect people in care homes (recognising that this was an emergency situation), the panel considered the lack of transparency around these wider uses to be a risk and questioned the linked data being used more widely without consent. It was suggested that people should be given the opportunity to provide consent for uses of the data that go beyond the original scope (in this case, helping understand the risk of and minimise the spread of COVID-19 in care homes).

“It could be used for good things in the future, but I don't think it's great you can take that system that exists for an emergency and then adapt it for future projects. If there was consent for the people in that household, there may be better awareness. But otherwise it feels quite intrusive.” (Session two)

The panel highlighted the importance of weighing up the benefits and harms that this use of data may have on individuals and society. While they could see the benefits of such data projects during the pandemic, there was also a sense of powerlessness in terms of how data about citizens were used. A clearly outlined public benefit for any future use of the CURL data was therefore deemed to be important.

“It's got to be the impact of the people and the communities. Why are they getting that data, and what would be the impact on the community? It's about having a clear purpose.” (Session two)

### Concerns about data security

Other concerns raised about this project were the amount of data being analysed, the extent to which personal data could be accessed via the CHI identification numbers, and the risk of data breaches occurring.

Given the possible future uses of CURL data, participants felt it was important to ensure adequate security was in place to prevent data leakages or misuse. They also felt that the amount of data being collected should be limited to minimise the impact on individuals if such incidents were to occur.

“The more organisations it's shared amongst the more susceptible it is to falling into the wrong hands. They've already mentioned they work with companies, so they know your age, your details. Am I going to be sold insurance products? Do they need all the data that's passed over to them? Are there safeguards in place for that, as well?” (Session two)

### Importance of data completeness

Participants also supported the idea of reviewing the data for any gaps that would risk individuals being excluded or not benefitting from initiatives if their CHI number was not known. This reflected a broader need for reassurance that data were being used ethically, robustly, and for the benefit of society. While it was agreed that the data should be as complete as possible, it was also felt that only the minimum amount of data required to fulfil the project objectives should be used.

“In theory, the more data there is, the more potential for misuse, or even use that wasn't its original intent”. (Session two)

### Wariness of private sector use of health data

Participants were wary of commercial interests in the CURL project. While reassured by the additional information provided by the specialist on the role of ethical committees in academia and in public bodies to control access to health data, participants felt it was important to consider the risk of misuse by private sector organisations, such as insurance companies. It was deemed appropriate that decisions about the use of health data should be made by the NHS.

In concluding discussions on the CURL project, the panel noted down their key ethical considerations on a digital whiteboard using post-it notes.

### Past project three: Equalities and protected characteristics

The equalities and protected characteristics project aimed to develop as complete a picture as possible of the protected characteristics across the Scottish population using information from medical records, education records and census data. The purpose of linking this data together was to enable public bodies and academic organisations to better consider equality issues when planning and delivering services. A summary of this project was presented to the panel in plenary by Duncan Buchanan (Research Data Scotland) and was followed by smaller breakout discussion and Q&A with Duncan in plenary.

### Benefits and risks of linked data for future use

The ability to quickly access this linked data in future was considered a benefit of the project. Reflecting on the pandemic, when data needed to be compiled or linked quickly, it was felt that having such data already available would ensure speed and quality if it was ever required urgently. A more exceptional view, however, was that this might result in having data “for the sake of it” and that a clear purpose was lacking.

“It's a good thing they've got access to data, especially following the pandemic so you can roll out help and things like that in a timely fashion and bring these bodies together.” (Session three)

Reflecting on the presentation, the panel felt assured that the organisations involved had been aware of the challenges associated with this type of data linkage and taken steps to address them. For instance, the panel were reassured about the existence of safe havens (secure

environments where data is held and can only be accessed by approved researchers). The panel also pointed to the use of various data sources to ensure the information being used was more accurate than if relying on only one source (like the census).

### Risk of incomplete data impacting decisions

The risks associated with this project hinged mainly on the issue of data quality. As had been pointed out in the presentation, one of the challenges with this project was accounting for the different ways in which characteristics were recorded and individuals' changing circumstances. For instance, there was no data available on gender reassignment or sexual orientation, and there were some characteristics (such as religion or disability) that were only recorded every ten years. There was some concern among participants that incomplete data could skew the results, providing inaccurate information and leading to "bad" policy decisions.

"They talked about gender reassignment surgeries not being tracked, but it made me think about broader data that might not be gathered. What you exclude can be very telling. If you don't take some data, or if people refuse to give it, then it still might skew results, and over time, that gets worse and worse". (Session three)

Other risks and challenges associated with this project included the possibility of identifying an individual due to the amount of information being collected across multiple sources; data being open to abuse if information was not stored securely or if passed onto third parties; and the lack of clear research objectives leading to data being used for purposes not in the public interest or that exacerbate inequality or discrimination.

### Importance of having a clear justification and set parameters for using the data

Given the possible future uses of this linked data, concerns were raised over data being passed to third parties and so it was felt that any organisation wishing to make use of this data would need a clear justification. There was a view that those seeking to use the data should demonstrate how this would benefit communities and be in the public good. Considering the reflections offered by the academic Anuj Puri on the projects presented in session three, the panel also highlighted the importance of staying within the original scope of a project, especially where the principles of consent apply.

"They're using people's information for a separate project where they haven't asked the individuals. It's been used for other things. I think people should be given the opportunity to say, 'We're going to give your data to a 3rd party,' and say yes or no. At the end of the day, it's for a completely different project." (Session three)

### The challenge of reconciling different ethical issues

These discussions highlighted a broader challenge for some participants in reconciling issues around identifiability, data quality, and consent. Participants still had questions about these aspects and how they related to each other:

### Some questions raised by the panel in relation to the equalities project:

- “To what extent does anonymity, or pseudonymisation, of data compromise data quality?”
- How do you minimise the risk to data accuracy while removing identifying information?”
- How do you get informed consent if the data is being anonymised?”
- “Can the ID number somehow be traced back to the individual/their information?”

Questions over the identifiability of data highlighted an ongoing lack of clarity around what impact pseudonymising or de-identifying data would have on other aspects like data quality and individual privacy. There was some reassurance in knowing that measures were in place to help remove some identifying information to protect individuals’ privacy. However, concern remained about the potential for data to become identifiable when linked in ways such as in the equalities project and participants suggested that the public may not be aware of this.

In concluding discussions on the equalities and protected characteristics project, the panel noted down their key ethical considerations on a digital whiteboard using post-it notes.

### Past project four: Ukrainian Displaced People

The UK and Scottish Governments processed and shared data during the Ukraine Crisis so that Ukrainians could be safely housed across Scotland. Immigration, safeguarding and housing data were shared between relevant agencies and organisations to ensure displaced peoples could be safely looked after. A summary of this project was presented to the panel in plenary by a Scottish Government representative and was followed by smaller breakout discussion and Q&A with the presenter in plenary.

#### Benefit of defining data use principles

The aim of the project, to support Ukrainian people coming to Scotland, was generally seen as a positive one. Based on the information given in the presentation, participants considered the principles established by those involved in the programme (such as data minimisation, necessity, proportionality, and humanitarianism) to be appropriate and felt that the use of data had been restricted in line with those principles.

“That's how things should be in general. Things should be defined at the very beginning, instead of collecting as much data as possible. Define the principles and then collect what you need.” (Session three)

#### Concerns about holding and updating data indefinitely

While recognising the humanitarian good of the Ukrainian Displaced People project, a number of risks and challenges were identified, such as holding sensitive information about Ukrainian people and their hosts for an indefinite period of time (given the uncertainty around Russia’s

invasion of Ukraine) and keeping the data up to date. It was also suggested that the process could be insensitive to Ukrainian refugees who had been through a traumatic experience.

"It seems like reducing people to numbers. Every family's journey has now become a case note." (Session three)

Discussions about this project reflected a broad range of views on wider debates around immigration and refugee policies. In weighing up the relative benefits and risks, participants had different groups in the forefront of their minds; some were thinking about the safety and wellbeing of refugees, and others that of the hosts.

### Concerns over data security in an international emergency

This project raised questions about data security, given the international context and the sharing of data between organisations during a period of conflict. The panel recognised the complexities and challenges around this but wanted to know more about how data were kept secure to protect the refugees and hosts, especially when such sensitive information (e.g. criminal records) was shared between different countries and agencies:

#### Questions raised by the panel in relation to the Ukrainian Displaced People project:

- "What differences are there between Scotland and other UK nations regarding the approach to data collection on Ukrainian refugees and hosts?"
- "How is data security managed on the Ukraine project?"
- "Who has access to data on hosts/refugees? Which delivery agencies/parts of the council?"
- "Is there any risk to 'group privacy' / a risk to the Ukrainian community from how data could be used?"

It was recognised that ensuring data is of good quality takes time, but in an urgent or emergency situation there was a risk that this could be overlooked. Having accurate information was considered important for avoiding any exploitation of individuals – both refugees and hosts – involved in the programme.

The panel raised further considerations in relation to how different countries might approach data sharing and retention, and how any potential differences are accounted for.

"The fact that you're looking at foreign nationals coming into the country. You're holding details about people from another country which needs to be held with sensitivity. You would presumably want to give that information back at some point and probably wouldn't want to hold onto it going forwards. One of the big concerns is, what do you do with the information going forward." (Session three)

In concluding discussions on the Ukrainian Displaced People project, the panel noted down their key ethical considerations on a digital whiteboard using post-it notes.

## Looking forward: possible future projects

This chapter summarises the panel's reflections on three projects that were in the early stages of development or that had not yet started, and that the DIN team was consulted on. These projects gave participants an opportunity to explore aspects of data use that they had not been introduced to before (such as use of emerging technologies, different data sources, and private or third sector involvement).

By exploring further reactions to and perceptions of emerging data projects with these new elements, this chapter highlights how the participants' tested and consolidated their thinking around the ethical principles that had been formed in the previous sessions, which were later developed into final guidelines.

The data projects reviewed in session five included:

- Little Knight – exploring how Artificial Intelligence (AI) could be used to identify patterns and correlations in anonymised school attendance records to safeguard children from abuse.
- Policing the pandemic – linking police and health data to understand the usefulness and fairness of the Coronavirus Regulations that were introduced during the pandemic.
- Mobility – using mobile phone app data collections to investigate how the flow of people in city centre workplaces or greenspaces changed during the pandemic.

An alternative perspective was provided in session five by Laura Carter (from the Ada Lovelace Institute), who offered insights on some key ethical considerations for these emerging projects.

Each project is presented separately, summarising the key ethical considerations raised by the panel in their initial review of each project. These reflections informed the ethical guidelines that the panel developed in the final session.

Despite the introduction of new approaches to using data (such as artificial intelligence, or AI), different types of data (such as police data and mobile phone app data), and different organisations (outside of the public sector), the ethical considerations raised by the panel were similar to those raised in the past project reviews.

## Key findings:

- In session five, participants reviewed new and emerging projects that the DIN team was considering supporting and discussed the key ethical considerations.
- On the **Little Knight** project, the panel felt there was a clear public benefit in terms of supporting social services and safeguarding children from abuse. There was some wariness around the use (and possible misuse) of AI.
- On the **Policing and Pandemic** project, the panel felt this was worthwhile and would help Police Scotland gather important insights as to how the pandemic affected people. Concerns were raised over the possible stigmatisation and increased surveillance of vulnerable groups.
- On the **Mobility** project, the panel considered the use of mobile data to be innovative and useful. It was also recognised that such data could lead to the removal of green spaces, but issues around consent were raised. “Is there any risk to ‘group privacy’ / a risk to the Ukrainian community from how data could be used?”

The **key ethical considerations** raised in relation to these new and emerging projects were similar to those highlighted in the past project reviews, and included:

- Weighing up the relative benefits and harms to groups in society.
- Proportionate use and not going beyond the original scope (including care over who data are shared with).
- Ensuring transparency and accountability in decisions about what data are used, by whom, and for how long it is held.
- Ensuring the principles of consent are adhered to.

## Little Knight

Little Knight is a not-for-profit initiative run through the Scottish Tech Army, which is a separate organisation bringing together specialist technical skills for charity projects. The Little Knight team are a group of volunteers from the private sector (with expertise in healthcare, education and social care) who came together to discuss how their technical skills could be used to help safeguard children from abuse. Little Knight was exploring how artificial intelligence (AI) could be used to identify patterns and correlations in anonymised school attendance records.

### Clear public benefit, but questions around AI

The potential public benefit of this research was clear to participants, as they felt it would help support social services and safeguard children from abuse.

There was some initial wariness around the use of AI in this project, with questions around how it would work in practice, what role it would play and what legal measures would be in place to regulate the use of such emerging technologies.

### Questions raised by the panel in relation to the Little Knight project:

- “What are the legal implications of using AI, where does the law fit in with AI?”
- “Would we be picking up the right information, the nuances implied by things that are written down, would that carry over from AI?”
- “If the AI’s learning from reports and flagging them up, and then a human would be involved to validate it, are there some [details] slipping through that should be flagged, and how would you know they’ve slipped through?”

After clarifications were provided by the specialist on how the AI element would work for this project, participants felt broadly assured that AI could be developed as a useful tool to support and speed up, but not replace, human decision-making and interventions.

“If we can rule out human error in some sort, that’s something we shouldn’t be afraid of.”  
(Session five)

The panel saw the value in using AI to support social workers in this way and felt that it could help reduce the risk of child abuse, however it was also deemed important to consider who it would benefit and how.

### Unintended consequences and misuse

While potential benefits were recognised, participants also felt that the use of AI posed some risks. It was felt that an overreliance on AI could lead to unintended consequences, such as social workers feeling judged on their performance and leaving the profession, or those abusing children trying to avoid detection through school attendance. Some participants described the prospect of trusting AI to do what a human does as “scary”, and there were concerns raised over the quality and robustness of the results it would produce. One particular concern was that it might unfairly target particular groups or miss some children at risk altogether.

“Because it’s artificial, it’s taking parts of the data and working with the most noticeable ones but it could also skip over data that seems normal, that seems fine. You’re getting children that attend school and from the outside everything looks perfect, but at the end of day you still need to make the calls, visit houses, do all the groundwork.” (Session five)

Aside from the use of AI, the scope of the project was another consideration that arose in the discussions, and participants felt that care would need to be taken over who data was shared with and for what purpose.

## Policing the pandemic

During the pandemic, the Coronavirus Regulations introduced unprecedented powers for UK police forces to ensure compliance in preventing the spread of COVID-19.<sup>14</sup> An Independent Advisory Group was set up during the pandemic to advise Police Scotland on the new powers and to ensure they were compliant with human rights. Police data – including records of encounters with the public during the pandemic and the database of Fixed Penalty Notices issued to members of the public under the Coronavirus Regulations – was used to inform recommendations on policing during the pandemic. The possibility of taking this research further was being considered, to understand the usefulness and fairness of the Coronavirus Regulations that were introduced during the pandemic. Police data would be linked with health data in Scotland’s National Safe Haven.

### Weighing up benefit and risk

The use of police data to understand the impact of the Coronavirus Regulations on people was considered worthwhile, with one participant describing it as “necessary”. There was some reassurance in knowing that these impacts were being explored and that Police Scotland was taking stock and reflecting on the groups that may have been adversely affected by the regulations.

“It’s a worthwhile cause to see how it did affect people in different areas.” (Session five)

However, the necessity of this project was not clear to all. Clarity was also sought over what exact data would be used, what it would be used for and whether individuals would be identifiable (which the specialist confirmed they would not be).

### Questions raised by the panel in relation to the Policing the Pandemic project:

- “Is it going to be patient or people-identifiable data you’ll be using?”
- “Would names be in the data, and if the names were in the data would they be there all the time?”
- “Does it still fulfil the original reason that it was put in place? It was acceptable at the time, the way we used the data. Is it the same thing we’re using it for, or different? Is there a time limit?”

---

<sup>14</sup> Link to Legislation.gov website: [The Health Protection \(Coronavirus\) Regulations 2020](#)

## Concern about the risk of stigmatisation

Participants were concerned by the possible future uses of the data outlined in the presentation, such as to explore whether underlying health-related vulnerabilities increased the likelihood of some individuals being subject to police enforcement for non-compliance. One breakout group questioned the assumption that health issues relate to non-compliance of the law. This prompted the specialist to explain that most incidents that come to the attention of the police are related to an underlying vulnerability such as mental health or addiction. Nevertheless, participants raised the possibility of stigmatisation of people with health issues and the potential for increased surveillance of vulnerable groups. There were also more general concerns raised, particularly among those with more sceptical views about the use of big data by governments, about the invasiveness of linking policing data with other datasets.

“I worry about the stigmatisation of people with health issues, and obviously as the professor said, it is a major issue, in policing, but whether or not it is the right approach, I do think there are lots of other reasons why lots of people in this situation would've broken the rules, that weren't health related. So it worries me there could be some form of stigmatisation of people with mental health or other health issues in this project.” (Session five)

While some were reassured by the specialist's clarifications about the purpose of the project (i.e. that the research is intended to understand the impact of enforcement on different groups in society), the panel emphasised the need for the public benefit to be defined and justified, with steps taken to minimise the potential harms to groups in society.

## Questions over scope and accountability

Participants were initially unclear on the relevance of combining health and police data for research and so felt that the purpose of a data project – especially when using such sensitive data – must be clear. They questioned whether the possible future uses of these data were within the original scope.

“If the data's being used beyond the agreed purpose, I'm not entirely sure about that one. It's gathered in a specific circumstance, and now is being moved to a different circumstance. I'm not sure that's right.” (Session five)

Accountability was also a key consideration, with participants highlighting the role of a bespoke public panel in the early stages of the policing the pandemic project as positive.

## Mobility

The Urban Big Data Centre (UBDC) is a research centre and national data service based at the University of Glasgow. UBDC promotes the use of big data and innovative research methods to improve social, economic and environmental well-being in cities. Their project would use mobile phone app data collections (anonymised and non-identifiable data) covering Glasgow City and neighbouring Council areas to investigate how the flow of people in places like city centre workplaces or greenspaces changed during the pandemic, both under lockdown and after the lifting of restrictions.

## Public benefit of using mobile data

Although there was some confusion over how mobile data were used, the project was thought to be innovative and potential benefits were identified. It was recognised that green space was beneficial to society and that the use of mobile data to understand how spaces are used could lead to better management of those spaces.

“It’s a great project, getting health improvements from green space is a really important issue, seeing how people utilise those spaces is also really, really good.” (Session five)

The public benefit was not clear to all participants, and there was a perceived risk that the data could be used to justify the closure of parks or reductions in green space. Ensuring that there is clarity of purpose and a justification for using the data to benefit society was therefore considered to be important.

## Consent for using mobile data

Mobile data was thought of as a “by-product” of smartphone usage and so its use was considered to be an effective way of gathering granular and accurate information that could be used for public good.

However, the issue of consent was raised in relation to this, with several breakout groups asking how the data were collected, and how consent was obtained by the private companies collecting mobile data.

### Questions raised by the panel in relation to the Mobility project:

- “Do you have to download this app, or is it data stored in the same place as where your health and public sector data is stored, or is it a different place than Glasgow?”
- “Is there a way to give consent or opt out?”
- “The companies using the apps, what sort of consent guidelines are in place with them, before the data’s even picked up?”
- “If you’re getting data from commercial bodies, do they have access to your outcomes?”

The specialist explained that mobile data were only available from those who had consented to it being shared, however the panel questioned the extent to which people would really know what they are consenting to. Although the panel were reassured by the processes outlined in the presentation to ensure the privacy, security and de-identification of mobile data, there remained some discomfort around the prospect of private sector organisations holding and selling mobile data without people’s knowledge. The panel felt it was important that people were given clearer guidance on this when it comes to the use of their mobile data.

“It’s interesting using mobile data and saying the legal basis is consent. A lot of people do turn on location data without thinking how that data is being used. I didn’t know there were companies out there that had the data.” (Session five)

# Ethical guidelines

## Themes underpinning the guidelines

The public dialogue raised a number of themes that were important for participants when considering past and future projects. These were reflected in the ethical guidelines formed by the panel and are summarised here for further context.

### Purpose

Participants felt that use of data by the public sector should have a clearly defined purpose and scope, to avoid misuse of individuals' data.

Throughout the public dialogue, participants expressed an interest in understanding the reasons for public sector use of data about citizens, how that data would be used, and who it would be used by. Participants generally trusted the public sector in Scotland to use their data appropriately, but felt that having these elements clearly defined and explained would provide the public with further reassurance that data were not being used inappropriately or without limit.

Underlying this need for reassurance was a broader concern about privacy, and a desire to avoid personal information about individuals being accessed unnecessarily. In the first session, questions such as “what information do they have about us?” and “why do they think they need it?” were raised, reflecting a need for greater understanding of the existing procedures in place for use of and protection of data. In line with previous public engagement research, participants were not fundamentally against their data being used, but wanted to know that it was being well-managed and protected.

“I am interested in the way data is used, and slightly concerned about the awful lot of data out there. It might be misused and there might be misinterpretation of data.” (Session one)

### Transparency

Participants wanted to see openness, honesty and clear information being made available about the use of data about citizens by the public sector.

Transparency has been an issue raised in previous public engagements on data use (as noted in the introduction) and this public dialogue was no exception. The need for transparency went hand-in-hand with the need for a clearly defined purpose and agreed scope. Participants felt that there should be openness about these aspects of a data-led project, with the public having the opportunity to find out more if they wanted to. Again, this came back to a desire for reassurance about data being well-managed and there being a degree of control over who can use their data.

“It comes back to that purpose. How clear are the reasons for [using data]? What will you then use it for? Is it just for that piece of work or is it being kept in use for lots of other things? It's all linked back to transparency. It's about being really clear about what it's for.” (Session six)

## Public benefit

The panel agreed that the use of data were only acceptable if there was a clear public benefit, or public good.

It is first worth noting that the terms public benefit, public good and public interest were used interchangeably by participants (though they mostly referred to public benefit). In session one, all three terms were introduced by one of the specialists as possible justifications for using data, but participants were not given detailed definitions and the distinctions between them were not outlined.

The concept of public benefit was difficult for participants to articulate. When asked how they interpreted this term, participants used fairly broad statements such as “something that benefits society”, “something that improves the lives of individuals” and contributes to “a happy society.” They felt that benefits could be wide-ranging, but suggested they would include improved quality of life, health and wellbeing, safety and security, and a reduction in harm. It was stressed, however, that public benefit or public good were subjective concepts, dependent on an individual’s values and ethics.

“Public good is never the same for everybody. What’s good for one might not be for another. It’s very difficult to tie down... I look at how it’s going to benefit me and my family.” (Session five)

There were differing views on whether public benefit applied to the whole population, or whether it could apply to a small section of society. On balance, after deliberation, it was felt that public benefit could apply to a smaller group as long as this was not at the detriment or harm of other people.

“Not every project will have benefits for every member of the public. That’ll never happen. But if it’s targeting minorities, it’ll definitely benefit those individuals, so it is still public good.” (Session five)

Public benefit and private sector involvement were not seen as mutually exclusive. While there were some concerns about private sector use of data, generally it was accepted that use of data by private sector organisations may contribute to a wider benefit to society. However, where the use of data were solely motivated by private sector profit, participants felt that this would not pass a public benefit test.

Ultimately, in developing their guidelines, participants did not give a specific interpretation of what public benefit means. They instead suggested that each data-led project should include an explanation of how public benefit has been interpreted and how the project meets that definition.

The issues relating to private sector involvement and public benefit were explored further in additional workshops. A separate report of the findings can be found on the Scottish Government website.

## Data quality

Participants stressed the need for decisions to be made on the best possible information available, meaning data should be as up to date and accurate as possible.

The quality of data was raised in the first session and continued to be a theme throughout the deliberations. Quality was linked with confidence in decision-making, with the feeling that data-driven decisions could only be reliably made if the data was accurate. This was a theme in participants' review of past projects. For example, they felt that inaccurate data might have led to people being left on the shielding list, or that gaps in equalities data might lead to policy decisions that do not adequately reflect the needs of those groups. Participants therefore wanted reassurance about quality assurance systems being in place, particularly if a data-led project involved multiple organisations that may have different quality standards.

“One aspect that stood out to me is the quality of data. You can have a mass quantity of data, but the usefulness could be zero. Quality over quantity.” (Session six)

There was also some scepticism about the quality of data being held by public sector bodies. For example, participants gave examples of receiving communications about their health conditions during COVID-19 which were not accurate. They therefore emphasised the importance of the public sector using up-to-date information, particularly when making decisions that could impact on health and wellbeing. This led to the suggestion that data sharing projects should be required to meet a minimum quality threshold before they go ahead.

Linked to data quality was a concern that data could be misinterpreted. It was felt that, even when data are up to date, it can be interpreted differently. While it was acknowledged that there may always be a risk of different interpretations of information, it was suggested that future data quality mechanisms should include some guidance on how to interpret data.

## Accountability

While there was recognition that legislation was in place to govern the use of data, participants felt that an independent body should oversee decisions about data use and hold organisations accountable for any misuse of data.

Much of the discussion around accountability related to the need for a system of governance and oversight of the public sector use of data. While participants were generally reassured by existing systems in place via GDPR and data protection legislation, they nonetheless stressed that organisations should be required to meet these standards and held accountable for any misuse. They therefore wanted assurance that there was oversight of organisations' adherence to these existing legislative conditions.

A clear process of governance and accountability was seen as particularly important when there were multiple organisations involved in sharing or using data, as there was a perception that data being passed between organisations would introduce a risk to data security. Having an independent body to oversee this process was therefore seen as an important way of minimising this risk. This led to the suggestion of an independent panel to help make decisions on whether public sector use of data should go ahead or not.

There was some confusion about the role of the DIN itself, and the extent to which it already fulfilled this oversight function. The information provided about the DIN gave some reassurance that there was a system in place to oversee public sector use of data, in particular the ethical framework that the DIN expects members to adhere to when running data projects. However, there was also confusion about what the exact role of the DIN was, and for some this persisted throughout the deliberations.

Overall, the emphasis was on a need for independent oversight of the use of data and clear sanctions for the misuse of data.

## Public involvement

Participants felt it was important for the public to be involved in decisions about public sector use of data but there were different views on the nature of this role.

One view was that the public should be involved in forming principles around acceptable use of data, much like the process used in this public dialogue, but that decisions about whether or not a project goes ahead should be confined to specialists with expertise in data. Another view was that the public may not be “data experts” but that they should still have a say on whether or not use of their data should go ahead.

Participant 1: “The public need to be there to say, ‘these are the objective principles I want you to adhere to for any [data] project.’ But when it comes down to judging a project on its merit, at that point you have to be confident that experts are in place to judge the project in line with the principles given by the public...”(Session six)

Participant 2: “At the end of the day, it’s public data so the public should decide if data is used for that project and every project should be the same.” (Session six)

## Urgency

There was an understanding that context can have an impact on whether data use is deemed acceptable or unacceptable. In an emergency situation, the panel felt that some flexibility in the guidelines surrounding data use may be needed.

When participants considered the different examples of public sector use of data, it became clear that there was no single, one-size-fits-all approach to deciding on whether and how data should be used. Participants acknowledged that during COVID-19, data needed to be accessed quickly and that it may not have been possible to consider all the potential ethical implications at the time. For instance, there was one view that it would not be appropriate to seek approval from a panel involving members of the public for the use of data in an emergency or urgent situation, as this would place an unreasonable burden on them to make decisions under pressure without time for them to understand the issues fully.

There was an understanding that, in an emergency situation where there is a threat to life, there may need to be flexibility in some of the ethical guidelines to ensure decisions can be made quickly by those in an appropriate position.

They therefore felt that each data-led project needed to be judged on a case-by-case basis. To help clarify whether guidelines might need to be flexed, participants felt it was important that there was an agreed definition of what constitutes an emergency. They felt this could refer to situations where there was a threat to life, but felt a more widely agreed definition was needed.

“It completely depends on the circumstances. An emergency, like COVID and the shielding list, is completely different to [a project about] the use of transport...we can't really treat it all the same.” (Session six)

### Guidelines produced by the panel

The core question that the panel was asked to consider was “What guidelines should the public sector follow when using citizens' data?” The guidelines developed by the panel are presented in this section, grouped under five key themes.

The guidelines were developed iteratively by participants over the course of the public dialogue. The wording of the guidelines largely came from participants themselves and reflect the language they used. Where any edits to wording were made by Ipsos, this was to correct repetition or duplication, to reorder points into a more logical flow, or to correct any minor points of fact.

#### When using citizen's data, the public sector should manage the scope by:

- Ensuring the purpose for using the data is clearly defined and the data is used only for that purpose. Timescales for use should be clearly defined.
- Having a clearly agreed justification for using citizen's data (i.e. if there is a clear public benefit) and ensuring that only data that is necessary for the project is used.
- Ensuring that data are not used solely (directly or indirectly) for profit by private sector organisations. The public sector should ensure that private sector partners only use data proportionate to the specific purpose it was collected for.
- Not using data outside the scope of any consent that applies to the data.
- Not sharing data beyond the agreed organisations. If more organisations are included later in a project, they should go through an ethical assessment.

#### When using citizens' data, the public sector should ensure there is transparency by:

- Making clear what data are being used and for what purpose.
- Making clear which organisations can access the data, and why.
- Specifying how long data will be stored for before deletion.

- Ensuring an ethical assessment is carried out once the scope of the project is known.
- Ensuring the public can easily access information about the project, including: what data are being used and for what purpose, how long data are stored before they are deleted, and a summary of findings or impact of project (where it is legally possible to do so and where individuals are not identified).

**The public sector should ensure the use of citizens' data is in the public benefit by:**

- Clearly defining and explaining what the public benefit is.
- Considering whether the public benefits of using the data clearly outweigh the risks. Any potential harms from use of the data need to be analysed and weighed against benefits.
- Considering negative impacts to the public and/or the environment or economy, with possible longer term impacts also considered. Projects that benefit or make a positive impact on a small number of people can be in the public benefit, provided they do not negatively impact others, the environment or the economy.
- Ensuring that identifiable data are only used if it meets the standard of achieving public benefit.

**When using citizens' data, the public sector should ensure data quality by:**

- Establishing and publishing a minimum quality standard for data projects (that includes consideration of how much data is needed). The extent to which data projects meet the threshold for data quality must be checked and continually assessed by the team delivering the project. If there is involvement from the private sector, these checks should be made by someone from government/public sector.
- Using up to date data that matches the agreed purpose and specific scope.
- Ensuring data are held securely for an agreed period after a project to allow for quality checking.
- Determining who can access the data and monitoring who has accessed the data.

**When using citizens' data, the public sector should ensure there is accountability by:**

- Clearly documenting the process used to decide whether the project should go ahead (to an agreed formal structure).
- Ensuring there is a hierarchical organisation chart to show who is responsible/accountable for each aspect/stage of the project.

- Seeking approval and oversight from an independent panel on whether a data project should go ahead or not, including whether public benefits outweigh risks. The panel should make decisions based on what is in the best interests of the public and there should be no declared conflicts of interest on the panel.
- Consulting members of the public on the acceptability of the use of the data (for determining principles but not to decide if a project should go ahead or not – this is the role of the independent panel).
- Ensuring an ethical assessment is carried out once the scope of the project is known.
- Taking responsibility when something goes wrong and stopping the project if necessary.
- Ensuring there is independent oversight from a third party (e.g. ICO and DIN) for projects involving the private sector, with clear sanctions for misuse (criminal and civil).

**When using citizens' data, the context should be considered, by:**

- Defining what constitutes an emergency. Any impacts of flexing guidelines in this context should be assessed continually, as far as practical, and after the fact (including any lessons learned).
- In an emergency situation, such as where there is threat to life, it may be necessary for data to be used that was not part of the original scope.
- In the event of an emergency the use of identifiable data can be justified. If the private sector is involved, there should be clear rules about what private sector organisations do with data after an emergency including when they are deleted.
- In an emergency situation, it may be necessary for the timescales for data retention and deletion to be reviewed and extended.

# Future engagement

One of the aims of this public dialogue was “to create a blueprint for a long-term, sustainable form for engaging and involving the public in data policy, scrutiny and decisions”. To help realise that aim, participants were asked what role the public should play in decisions about public sector use of data, and specifically how the public should be engaged on this topic in future. This section outlines views on these topics, as well as participants’ reflections on their own experiences as part of the panel.

## Role of the public in decisions about public sector data use

Participants felt that the public had an important role to play in helping shape how data were used by the public sector. They suggested that members of the public could provide a balance to the views of experts and data specialists, potentially offering new ideas or alternative issues to consider.

“Experts are, for me, looking at one thing and one thing only. The public have got their own perceptions and can input more feeling and reality into this.” (Session six)

If the data in question had originated from members of the public, it was seen as only fair for the public to have a say in how the data would ultimately be used.

“It’s always beneficial and interesting to have an outsider’s perspective...it’s our data, and I think we should have a say on how it’s dealt with.” (Session five)

It was also suggested that the public could help to ensure use of data is explained in “lay person” terms, by asking questions and encouraging data specialists to clarify what might otherwise be very technical information. Participants considered themselves to have played this role in the panel.

Involvement of the public was seen as a marker of transparency, one of the key factors that participants felt was important for building trust in public sector use of data. As previously noted, there was a generally high level of trust in Scottish Government and the public sector to begin with. Over the course of deliberation, this feeling of trust had either remained high, or had increased. Participants attributed this to their own process of learning as a panel, specifically in relation to the current data protection landscape and the steps the public sector had to go through before it uses data.

“Starting out, I had a very low [awareness] of how data was being used. Over the course of this panel, it became a lot less nebulous...it has increased my confidence, as it has showed how much vigour is involved in getting ethics through... I do trust data usage more.” (Session five)

In contrast, when decisions were made by the public sector without any involvement of the public, they felt this created a sense of distrust in those decisions. One participant used the

example of changes to the organ donation systems in Scotland, which they thought had been transparent because of the level of public involvement and communication around it.

There was an acceptance that it may not be possible to seek the views of the public on data-led projects in an emergency situation. However, when this was the case, it was felt that the public should at least be informed about how data were being used.

### **How the public should be engaged in future**

There was overwhelming support for future public engagement on the use of data, which reflected participants' positive feelings about their own experiences as members of the panel.

Echoing earlier views about the importance of transparency, participants felt that data-led projects should be widely publicised and promoted to help raise public awareness and interest.

“More publication and promotion of [data projects] would be exciting for the public...and that knowledge could add to public scrutiny.” (Session five)

Participants viewed a public panel, designed and structured in a similar way to the one they were part of, as a good way of engaging the public. They felt they had benefitted from having the opportunity to learn about the topic, hear from and speak with the experts, and then reach informed conclusions. This learning process was seen as particularly important when asking the public for views on technical, complex topics such as the use of data.

It was suggested that an ongoing panel could be used to help make decisions about future use of data by the public sector. Potential uses for a panel could be to review and provide feedback on potential data-led projects, or to revisit ethical guidelines developed by this panel to test whether they were still appropriate.

“Have a pool of people that they could draw on, who are interested...[to provide] those checks and balances... to challenge some things. I'm not entirely sure how it would look, but I think it's a good idea having lay people on these [panels] in some form.” (Session six)

Other suggested forms of engagement included teaching children about data at school, and using websites (such as a public sector website) to invite feedback from the public about potential data-led projects. However, it was felt that online consultations might only appeal to a certain type of person, and that a randomly chosen sample of the public, like the approach used for this panel, would help to ensure involvement from more diverse groups in society.

### **Reflections on participants' involvement in this panel**

This public dialogue supported participants to express a range of views on different types of data projects, and explored their expectations and understanding of the ethical considerations for future use of data about citizens. The panel's reflections in the final session highlighted that learning journey, with participants describing how they went from feeling “overwhelmed” in the first session to feeling “informed” and “empowered” by the end. The image below shows the participants' experience of the deliberative journey from the first to the last session, expressed in their own words:

**Figure 1.3: 3 words to describe the session (online community feedback)**



One positive aspect of the process was the opportunity for participants to meet (virtually) and engage with each other, particularly through the smaller group discussions. They felt that the process had helped them to realise their own biases, and to listen to and be shaped by each other’s views.

It was clear that the deliberative nature of the public dialogue had been beneficial for participants. They welcomed the opportunity to learn about the topic in depth, hear from and speak with a range of experts, and then reach informed conclusions. They also appreciated the ability to have direct engagement with expert specialists, made possible by bringing them the expert speakers into the smaller group discussions. It was common for participants to reflect on their overall learning journey, and the growth in their own understanding about the use of data.

“In the first week, I think it felt very overwhelming.... it did become clearer as we went through the weeks...it's nice to be a part of something that you think you might have a slight impact on something moving forward.” (Session six)

The main drawback they highlighted about the process was that the information in the early stages was overwhelming, and this made some feel confused and “out of their depth”. For future public dialogues or similar forms of engagement, participants suggested that information provided in the early stages of the process should be as simple as possible, and that dense presentations of technical information should be avoided.

Overall, there was a sense that involvement in the panel was worthwhile and that they were genuinely having an impact on future policy.

"It really feels like you're actually connected with a process that will change not just your life, but the lives of other people." (Session six)

There are aspects of the process which may, in and of themselves, have impacted on how participants engaged with the topics, such as:

- **The delivery of presentations:** while a template for presentations was provided to ensure key details were covered, specialists interpreted these details in different ways in relation to their projects, and the variations in delivery style may have influenced how participants responded to, and engaged with, the projects. Where participants felt presentations used too many 'academic' terms or were too 'jargon'-heavy, they found the discussion afterwards more challenging. Where presentations were felt to be clearer and succinct, it was easier to get straight into the discussion afterwards. Ongoing opportunities for Q&A with specialists helped participants clarify their understanding. Having specialists available to join breakout discussions was also beneficial in addressing any issues or misunderstandings, allowing participants to progress their discussions.
- **Engagement approaches:** some presentations were delivered live (or were pre-recorded and played back) during plenary, while others were delivered directly to participants in small breakout groups. Feedback suggested that this worked well in terms of getting the specialists closer to participants and enabling direct feedback and Q&A, but also resulted in less time for participants to reflect on and discuss the project. The specialists being present during discussions may have resulted in participants being less willing to share their views.
- **Perceived complexity of the project:** some data projects, such as those related to the pandemic, resonated more with participants' own lived experiences while others, such as the data linkage projects, felt more abstract and this may have impacted on how participants responded to them.
- **Variance in online community engagement:** the online community was primarily a vehicle for maintaining engagement with the panel in-between sessions, however any data collected from it (such as survey tracking) has been treated with caution as not all participants chose to join and, of those who did, not all activities were completed. While most tasks via the online community were discrete and did not inform the main panel process, participants who did not register on the online community were offered opportunities to participate in certain tasks (such as voting on projects they wanted to hear about) via email instead.

# Conclusion

This public dialogue set out to:

- Explore the ethical implications arising from the use of data by the public sector.
- Develop a set of principles to inform future data-driven research projects and policies.
- Explore the possibility of a longer-term approach to public engagement as a means of providing external scrutiny of the public sector.

This report has highlighted the thoughtful ways in which participants engaged with these big issues and the learning journey they went on to develop their final ethical guidelines.

In this concluding chapter we outline the key findings from this public dialogue that add to the existing body of knowledge on public attitudes to data. It also provides reflections on public dialogue format and the potential for future public engagement on data use.

**The panel were generally trusting of the public sector's use of data, but wanted reassurance that ethical principles would be followed alongside existing legal frameworks.**

From the start of the process, participants had fairly high levels of trust in the Scottish Government and wider public sector. This trust was linked to an understanding that use of data by the public sector can have benefits for society, particularly in relation to health emergencies such as COVID-19. Participants were more inclined to trust the public sector with their data than the private sector, as they felt the was more likely to be motivated by public good rather than by commercial gain.

Views on the role of the private sector did develop over time. Although in earlier sessions there was scepticism over the trustworthiness of the private sector, by the final session there were more nuanced perspectives. One perception was that some private sector companies can provide benefit by using data and that it could be acceptable as long as their use of data was not solely driven by profit, or to only benefit an organisation and not the wider public. However, there were differing views on this, with some feeling that involvement of the private sector in any public sector data projects was problematic. As a result, the final guidelines relating to private sector use of data about citizens were not endorsed by all. Despite not being a focus of the panel, participants said they would have welcomed more space and time to explore the role of the private sector in detail. This issue was explored in additional workshops (the findings of this can be found in a separate report on the Scottish Government website).

This sense of trust was based on an expectation that the public sector would follow rules and regulations around the use of data and be held accountable for any misuse. Having learned more about existing data protection legislation, the panel generally felt reassured that systems were in place to govern the use of their data. However, the existence of data protection legislation did not remove the need for ethical principles to be followed and it was clear that both

legal and ethical considerations were important for building public trust. Though some of the panel's ethical guidelines reflected the content of existing legislation such as GDPR (for example principles of purpose, transparency, and accountability), participants nonetheless viewed these as fundamental ethical considerations for the public sector to follow.

**Views were influenced by the context in which data was being used, specifically whether it was an emergency situation or not. The panel therefore identified a need for some flexibility to be allowed for, while adhering to basic principles.**

A unique feature of this public dialogue was that participants were given the opportunity to scrutinise and share feedback on real public sector data projects. As some of those projects had been delivered under the unprecedented circumstances of the COVID-19 pandemic, this raised specific issues for the panel around how the public sector should approach data in an emergency.

The panel understood that data had to be used quickly to support decisions related to COVID-19 and they recognised the benefits of doing so. To take the shielding list as an example, it was acknowledged that the Scottish Government and partners were dealing with exceptional circumstances and that use of data had public health implications. The panel's ethical guidelines therefore reflected the need for agility, allowing for quick action to be taken in the event of another pandemic or similar emergency. As a minimum, however, there was an expectation of transparency around what data were being used and for what purpose.

**The panel emphasised the importance of data quality as an ethical consideration when using data. They wanted to see a minimum quality standard put in place which future data sharing projects would be required to meet.**

In this public dialogue, the panel discussed data quality in some detail and this became an important factor as they developed their final ethical guidelines.

Data quality was seen as a measure of accuracy, and the panel stressed the importance of data-driven decisions being made with the most up-to-date and accurate information. Underlying this view was a recognition of the impacts of some of the decisions that the public sector had, and would have, to make through their use of data. The obvious example, again, was the shielding list and the importance of this being based on accurate information.

Data quality was also linked to fairness, with the panel feeling that gaps in data could lead to individuals being excluded or not benefitting from certain initiatives. The panel also recognised that the quality of data may differ between different groups in society, and they questioned whether this might impact on policy decisions.

The importance of data quality was reflected in the panel's suggestion of a minimum quality standard for each data project to meet. This recommendation applied to all projects, but where there was private sector involvement the panel specifically wanted to see checks on data quality carried out by the public sector. This, again, reflected the higher levels of trust the panel placed in public sector organisations in comparison with the private sector.

### **The panel was sensitive to the impacts of data use, both positive and negative, on marginalised groups.**

The potential use of special category, or sensitive, data raised specific ethical considerations for the panel. Protection of individual privacy was seen as particularly important when dealing with sensitive data, and the panel had concerns that misuse of this type of data could lead to individuals being discriminated against. For example, when reviewing the Policing the pandemic project there was concern that the use of data on health-related vulnerabilities might increase the likelihood of individuals being subject to police enforcement for non-compliance. Participants raised concerns about the risks of vulnerable groups being stigmatised on account of characteristics such as their underlying health conditions.

The panel's sensitivity to the impact on marginalised groups was also reflected in their discussions about public benefit. As they grappled with their own interpretations of public benefit, they arrived at the view that this can mean benefits to society as a whole but in some cases can also be restricted to a specific, smaller group. In the latter scenario, they felt there can be public benefit as long as this was not at the detriment or harm of other people.

### **The process highlighted the value of informed dialogue, and that the opportunity to learn about and deliberate on topic can lead to more informed decision making.**

This public dialogue has demonstrated the ability of the public to engage with a complex, and in some cases unfamiliar, topic and develop thoughtful principles for the future. The opportunity to have dialogue with experts, and with each other, helped participants to formulate their views on the most important ethical considerations.

However, as noted in the previous chapter, the panel's response to certain topics may have been influenced by presentation style and delivery. Over the course of the public dialogue, participants heard from twelve different experts and it was inevitable that these would resonate differently with participants. This emphasises the importance of introducing content gradually and in different formats, giving participants the space to ask questions and seek clarification on key points, and being open to different ways of presenting the information.

### **There was a great deal of support for future public engagement on the use of data by the public sector, and a public panel was seen as a good approach.**

The panel strongly believed that the public had a role to play in decisions about the use of data by the public sector. A unique aspect of this public dialogue was that participants acted as a panel, meaning they had the opportunity to review and appraise past, current and future data-led projects. The panel-style approach placed participants in the role of evaluator, giving feedback (sometimes directly to those involved) on data-led projects in a way that could influence decisions around their future delivery. The panel-style approach (with the group meeting over a four month period, with gaps of up to two weeks between workshops) also meant that they had a fairly long period of time to immerse themselves in the topic, reflect in between sessions and gradually develop their ethical guidelines in response to what they learned.

Reflecting their positive views on their own experience, participants felt that a public panel was a good model to replicate. They specifically felt that the opportunity to hear from and engage with experts, and to provide feedback on specific projects, were valuable. There was also clear appetite for future public engagement to feedback on specific data-led projects.

# Appendices

## Appendix A: Previous public engagement in Scotland

This includes workshops, citizens' assemblies, citizens' juries, focus groups, qualitative interviews, surveys, discrete choice experiments, game-based approaches, ethnography, public panels and consumer panels. The key insights from these public engagements fed into the design of this public dialogue and are summarised below.

Previous public engagements found general support for data sharing, across different contexts and through the different methods used to involve the public. There was also broad support for greater involvement of the public to inform government decisions and policy. Some engagement work highlighted particular conditions that should be met regarding the context and scope of data use that would increase participants' willingness to share data, with concepts such as public benefit being used. Yet, previous engagement also revealed complexities around participant views of 'public benefit' and 'public good' in Scotland. Some case studies revealed wide-ranging ideas of public benefit, with participants unwilling to clearly define public benefit in a way that might exclude some uses of data that they considered valuable. For example, deliberative workshops conducted in 2017 on the topic of what is meant by 'public benefit' found that participants' preference was for the widest possible public benefit to be felt by all, but they also acknowledged the value in research aiming to primarily benefit vulnerable groups within society.<sup>15</sup>

Transparency was regarded as desirable, but the term was used and understood differently by different engaged publics. Authors of a report on a series of 2010-2011 focus groups and workshops on trust and trustworthiness in data use concluded that transparency could include: 'informational transparency requiring disclosure of information on which decisions are based; participatory transparency, enabling public participation in decision-making processes; or, accountability transparency whereby decision-makers are held accountable'.<sup>16</sup>

Early engagement work that considered data governance did not reveal any strong preferences for particular approaches towards oversight, assessment, and accountability. Later engagement work also revealed some ambiguity regarding public participation in data governance. More active public participation was often seen as desirable. However, not all public engagements revealed a desire for greater citizen control. Some engagements found, for example, that not

---

<sup>15</sup> Aitken, McAteer, Davidson, Frostick, and Cunningham-Burley. 2018. 'Public Preferences Regarding Data Linkage for Health Research: A Discrete Choice Experiment'. *International Journal of Population Data Science* 3 (1): 429. <https://doi.org/10.23889/ijpds.v3i1.429>.

<sup>16</sup> Aitken, Cunningham-Burley, and Pagliari. 2016. 'Moving from Trust to Trustworthiness: Experiences of Public Engagement in the Scottish Health Informatics Programme'. *Science & Public Policy* 43 (5): 713–23. <https://doi.org/10.1093/scipol/scv075>

everyone had the requisite expertise to contribute to specific decisions and that family and work commitments would prevent people from engaging.<sup>17</sup>

Some insights have been gained on the complexities of mechanisms of consent and, to a lesser extent, on data quality. However, a contentious topic in data use is the involvement of the private sector, with higher trust being given to public bodies such as the NHS. As found in a 2022 deliberative workshop conducted by Ipsos Scotland on behalf of DataLoch, there was also some acceptance of private sector involvement, with conditions, such as exchanging data sharing for benefit.<sup>18</sup>

## Appendix B: meetings and members of the oversight group

The oversight group met four times over the course of the panel to advise on different stages of the project:

- **August 2022:** discussion and approval of the methodological design
- **October 2022:** discussion around early fieldwork progression
- **November 2022:** discussion around later fieldwork progression
- **January 2023:** discussion of the guidelines formulated by the panel and possible topics for further engagement

The members of the oversight group were:

- **Scottish Government representatives**
- **Academia:**
  - Prof. Rowan Cruft (University of Stirling)
  - Dr. SJ Bennett (University of Edinburgh)
  - Dr. Fay Niker (University of Stirling)
- **Civil Society:**

---

<sup>17</sup> Davidson, Sara, Christopher Mclean, Steven Treanor, Mhairi Aitken, Sarah Cunningham-Burley, Graeme Laurie, Claudia Pagliari, and Nayha Sethi. 2013. 'Public Acceptability of Data Sharing Between the Public, Private and Third Sectors for Research Purposes'. Gov.Scot. 2013. <http://www.gov.scot/publications/public-acceptability-data-sharing-between-public-private-third-sectors-research-purposes/>

<sup>18</sup> Ipsos. 2022. 'Public Perspectives on Access to Health Data by Non-Traditional Researchers: Findings from Deliberative Workshops'. IPSOS Scotland. Accessed 14 September 2022.

file:///Users/jamiewebb/Library/Mobile%20Documents/com~apple~CloudDocs/Public%20deliberations%20on%20access%20to%20health%20data%20by%20non-traditional%20researchers.pdf

- Representatives from the:
  - Ada Lovelace Institute
  - DemSoc Edinburgh
  - Urban Big Data Centre, University of Glasgow
- Chris Mackie (The ALLIANCE)
- Mariano Delli Santi (Open Rights Group)
- Prof. Roger Halliday (Research Data Scotland)
- Shayda Kashef (Administrative Data Research UK)
- Stephen Peacock (Information Commissioner's Office)

### Appendix C: sampling and recruitment

The [Sortition Foundation](#), a recruitment organisation specialising in representative random sampling, conducted recruitment for the public panel by sending 6,000 invitation letters across Scotland, using the Royal Mail Postcode Address File. Those living in more deprived areas were over-sampled to account for the lower response rates that are typically found in these areas. Recipients of the invite letter were signposted to an online form to register their interest.

Based on all those who registered their interest in joining the panel (293), a randomised stratified selection process took place that broadly reflected the demographics of Scotland, including age, gender, region, ethnicity, disability and education. Ethnic minority groups were over-sampled at the selection stage to ensure sufficient representation of these groups. An attitudinal measure was also included in the selection process to ensure a range of views were represented in terms of trust in the Scottish Government and public sector agencies to use data for the public good.

Overall, 30 people were selected to join the panel and 25 participated throughout. A table summarising the demographic profile of the final selected and confirmed sample can be found below.

**Table 1.1: Demographic profile of panel**

	Quota group	% in population	Number selected for panel
<b>Age</b>	16-24	11%	3
	25-34	18%	6
	35-54	32%	10
	55+	38%	11
<b>Gender</b>	Woman	52%	16

	Quota group	% in population	Number selected for panel
<b>Region</b>	Man	48%	13
	Non-binary/other	<i>No clear data</i>	1
	Central	12%	3
	Glasgow	13%	4
	Highlands and islands	8%	2
	Lothians	15%	5
	Mid Scotland and Fife	12%	4
	North East Scotland	14%	5
	South	13%	4
	West	13%	3
<b>Ethnicity</b>	African, Caribbean, Black or Black Scottish/British	1%	2
	Asian, Asian Scottish or Asian British	3%	3
	White Scottish/Other British/White Other	96%	24
	Other ethnic group or mixed/multiple ethnic groups	0%	1
<b>Disability</b>	No long-term physical or mental health condition	70%	22
	Long-term physical or mental health condition which is limiting	24%	6
	Long-term physical or mental health condition which is not limiting	6%	2
<b>Education</b>	Level 4 (Degree, Professional Qualification)	32%	10
	Level 3 (HNC/HND or equivalent)	13%	4
	Level 2 (Higher, A level or equivalent)	17%	5

	Quota group	% in population	Number selected for panel
<b>To what extent would you trust the Scottish Government and public sector agencies (for example councils or health boards) to use your data for the public good?</b>	Level 1 (O Grade, Standard Grade or equivalent)	17%	7
	No qualifications	15%	4
	A great deal	<i>No data</i>	7
	A fair amount	<i>No data</i>	14
	Not very much	<i>No data</i>	2
	Not at all	<i>No data</i>	2
	It depends on the department or agency	<i>No data</i>	3
	Don't know	<i>No data</i>	2

## Appendix D: methodology

### Overview of process

During the learning phase, participants heard presentations introducing them to the DIN, data protection and data ethics. In session two and three they heard presentations from representatives of the past projects the DIN were involved in, and in session five from individuals involved in potential future data projects. In sessions three and five, they also heard from independent academics who offered more of an outside perspective on the ethical issues related to the data projects.

After each presentation, participants moved into small breakout groups to discuss and reflect on what they had heard and share their thoughts. In the breakout discussions, participants agreed on clarification questions which were then answered by the speakers in the main plenary, or via a Q&A document which was shared with participants on an ongoing basis with written responses provided by the speakers, the Scottish Government and Ipsos.

Sessions three to six each began with the chair reflecting on what participants had discussed in their groups in the previous workshops. This provided a space for participants to reflect on where they had got to. At points throughout the workshops, each facilitator would provide 'flavour' feedback on their group's discussion so that participants had the opportunity to hear from others. Based on rapid analysis of the discussions by the research team, and reviewed by facilitators of breakout room discussions, the final sessions provided the panel with draft guidelines for review and ratification in breakout rooms. The rapid analysis has since been validated with systematic analysis, which was conducted post-fieldwork to inform this report.

**Table 1.2: Session summaries**

	<b>Date and time</b>	<b>Objective</b>	<b>Session description</b>	<b>Presentations and speakers</b>
<b>Session 1</b>	Tuesday 27 September 18.00-21.00	Introduce participants to the process, aims and role of the DIN	Panel was introduced to each other and familiarised with the process and topic area. Participants shared initial views and perceptions on data and how it is used by Scottish Government and public sector agencies, learned about the role of the DIN, data ethics and the legal context of data use.	DIN member: introduction to the Network  Nayha Sethi (UoE): introduction to data ethics  Stephen Peacock (ICO): introduction to data protection  Presentations delivered in plenary and followed by small breakout discussions.
<b>Session 2</b>	Saturday 8 October 10.00-13.00	“Looking back” part 1 – reviewing past projects related to COVID-19 pandemic	Panel developed an understanding of the types of projects that the DIN have delivered and evaluate past projects relating to the COVID-19 pandemic, considering the implications of using data about citizens in such circumstances.	Scottish Government: shielding list project summary  Dave Grzybowski: CURL project summary  Presentations delivered in plenary and followed by small breakout discussions.
<b>Session 3</b>	Tuesday 25 October 18.00-21.00	“Looking back” part 2 – reviewing past projects not related to COVID-19	Panel continued to develop their understanding of the types of projects that the DIN delivered and evaluate past projects, considering the ethical implications of using data about citizens in different circumstances. Participants also heard an outside perspective on the projects from Dr Anuj Puri, who shared his independent reflections on the ethical risks and challenges.	Duncan Buchanan (Research Data Scotland): equalities and protected characteristics project summary  Scottish Government: Ukrainian Displaced People project summary  Dr Anuj Puri (Tilburg Institute for Law, Technology and Society): An outside perspective on the data projects

	Date and time	Objective	Session description	Presentations and speakers
<b>Session 4</b>				Presentations delivered in plenary and followed by small breakout discussions.
	Saturday 12 November 10.00-13.00	Forming draft principles	Panel continued to discuss and consider ethical implications, starting to form principles that they think should apply to data projects.	None
<b>Session 5</b>				Ellen Ward (Scottish Tech Army): Little Knight project summary  Susan McVie (UoE): Policing the Pandemic project summary  Michael Sinclair (UoG): Mobility project summary
	Thursday 17 November 18.00-21.00	“Looking forward” – consider possible emerging and future projects	Panellists learned about emerging and future data projects and test draft principles. Participants also heard an outside perspective on the projects from Laura Carter, who shared her independent reflections on the ethical risks and challenges.	Above presentations delivered to small breakout groups (in a carousel format) with the speaker available for immediate Q&A and discussion. Below presentation delivered in plenary and followed by small breakout discussions.  Laura Carter (Ada Lovelace Institute): An outside perspective on the data projects.
<b>Session 6</b>	Saturday 3 December 10.00-13.00	Forming final guidelines	The panel finalised a set of ethical guidelines that they think should be followed by the Scottish Government and public sector when using data about citizens.	None

**Analysis**

The deliberative nature of the project allowed for ongoing analysis throughout fieldwork, meaning the research team observed discussions and checked in participants during each workshop to ensure key concepts (including content presented by specialists) were understood, to identify any areas where further clarification was needed, and establish emerging themes. The facilitation team debriefed after each session and discussed findings. This ensured that emerging principles and themes - both from workshop discussions and online community activities - could be played back to participants as the dialogue progressed. Each step in the analysis involved:

- **Note-taking:** there were live note-takers present at each session and in each breakout group to ensure conversations were recorded accurately. With participants' permission, each breakout group was audio recorded as a further record of discussion.
- **Debriefing:** facilitators came together the day after the workshop to share key themes and reflections from their group's discussions. The discussions were typically structured around the topic guide (and activities from the online community that week) and the core research team chaired these sessions to ensure all aspects of the session and online activities were covered. Facilitators drew on their own notes as well as the full transcripts.
- **Developing themes:** early findings from the workshops and online community were condensed into key points that were played back to the panel via a presentation summary delivered by the chair at the beginning of each workshop. In the following breakout discussion, participants were given the opportunity to reflect on these points and confirm/challenge those that did or didn't resonate with their experiences. This also provided an opportunity to check participants were clear on things, identify themes that resonated most strongly, and unearth any outstanding issues.
- **Ongoing analysis:** to support ongoing analysis, a spreadsheet was developed and updated as fieldwork progressed. A separate tab was created for each session, with the columns covering each breakout and key discussion questions and the rows summarising each group's discussion. The facilitators completed the analysis spreadsheet after each session, drawing on the full transcripts, recordings and their own notes. A summary column at the end of each tab enabled facilitators to note down key emerging themes and reflections.

## Appendix E: Q&A document

### About this document

As part of the panel process, members have the opportunity to ask the speakers questions about their presentations. Some questions are addressed during the session and any remaining

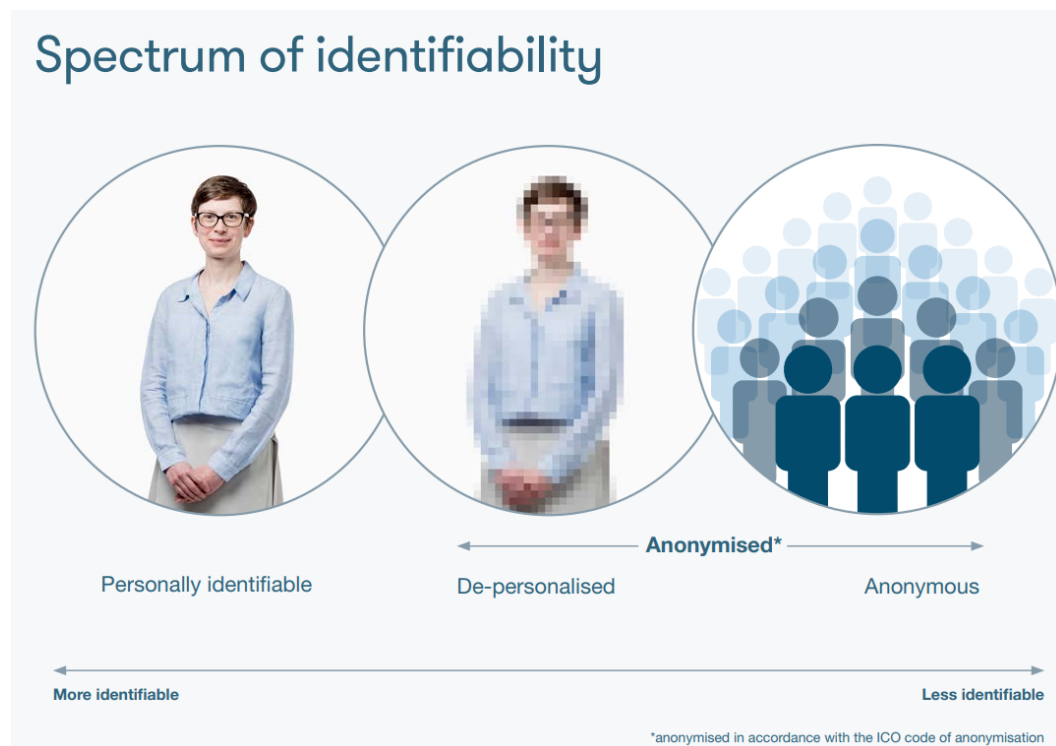
questions are collated in this Q&A document which is shared with the relevant speakers and the Data & Intelligence Network for response. A glossary of terms is also provided to help with some of the more technical language used throughout the panel.

The questions are organised by session and by theme. The document is available to panel members at any stage of the process.

## **Glossary of terms**

**Algorithm governance** - the use of algorithms and artificial intelligence in governance. Understanding the social implications of artificial intelligence, big data, and automated decision-making on society gives rise to concerns of transparency and accountability, among other ethical issues.

**Anonymisation of data** – the process used to prevent someone's personal identity from being revealed in a given set of data. The technical language of identifiability is complex. Many different words are used to describe the same thing, and many of those words are unnecessarily technical (for example pseudonymised, key-coded, de-identified for limited disclosure). It is important to explain clearly what it means when information is 'anonymised' and what the likelihood of re-identification is when using different types of data. The picture below tries to explain this, and you can read more about it [here](#).



(Source: Understanding Patient Data)

**Communities of practice** - groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly. In this case, the Network supports communities of practice on the ethical use of citizens' data.

**Constitutional privacy** – or “decisional privacy” refers to the freedom to make one’s own decisions without interference by others in regard to matters seen as intimate and personal. This is linked with informational privacy (see below).

**Contextual integrity** - requires that information gathering and sharing be appropriate to that context. An example of this can be restrictions on communication of patient information outside the healthcare context.

**Contestation** – refers to the ability to challenge outcomes determined by automated processes in governance. This is necessary to protect rights, ensure accountability and enhance public trust

**Data ethics** - the benefits, risks and wider social harms that should be considered when thinking about how data is used, such as when used by Network members for different types of projects.

**Data justice** – fairness in how people are treated, represented, and ‘seen’ by virtue of data processing. In large scale data use it is important to consider how that data might lead to bias or discrimination against groups of people.

**Data & Intelligence Network** – The DIN is a collaboration made up of members across the Scottish Public and Not-for-Profit Sectors, including Health Boards/Agencies, Local Authorities, Academia and third sector. The DIN is led by dedicated team of Scottish Government staff who actively work with members of the network to help deliver projects and support communities of practice. Throughout the panel it might be referred to as the DIN or the Network.

**Data Protection Act** – controls how personal information can be used and your rights to ask for information about yourself. Under the Data Protection Act 2018, you have the right to find out what information the government and other organisations store about you. Together with the UK GDPR, this forms the UK’s data protection legal framework.

**Data Protection Impact Assessment (DPIA)** - a process to help organisations to identify and minimise the data protection risks of a project.

**Data Controller** – the data controller determines the purposes for which and the manner in which personal data is processed. It can do this either on its own or jointly or in common with other organisations. This means that the data controller exercises overall control over the 'why' and the 'how' of a data processing activity.

**Data minimisation** - requires that the collection of personal information be limited to what is directly relevant and necessary to accomplish a specified purpose. Data should also be retained only for as long as is necessary to fulfil that purpose.

**Data Processor** - act on behalf of, and only on the instructions of, the relevant controller.

**“Five Safes”** - a set of principles which enable data services to provide safe research access to data. You can read more about the five safes [here](#). The principles are:

- **Safe data:** data is treated to protect any confidentiality concerns.
- **Safe projects:** research projects are approved by data owners for the public good.
- **Safe people:** researchers are trained and authorised to use data safely.
- **Safe settings:** a SecureLab environment prevents unauthorised use.
- **Safe outputs:** screened and approved outputs that are non-disclosive.

**Information privacy** – the control over one’s information because an individual’s choices can be influenced on the basis of information about them or others like them.

**Informed consent** - under the General Data Protection Regulation (GDPR), “for consent to be informed and specific, the data subject must at least be notified about the controller’s identity, what kind of data will be processed, how it will be used and the purpose of the processing operations.” From a privacy perspective, principles of informed consent require that the consent must be unambiguous, specific, informed and freely given.

**Group privacy** - refers to the collective interest in privacy and is concerned with the use of information and inferences drawn at a group rather than individual level. Collective interest in privacy arises out of the use of information concerning one member of a group to undermine the autonomy of other members of that group.

**UK GDPR** – this is our version of the EU’s General Data Protection Regulation (GDPR) and it controls how your personal information is used by organisations, businesses or the government. Together with the Data Protection Act, this forms the UK’s data protection legal framework.

**Non-anonymised data** - data which does contain identifiable information (e.g. name, address).

**Precautionary principle** – has its origin in Environment law and “enables decision-makers to adopt precautionary measures when scientific evidence about an environmental or human health hazard is uncertain and the stakes are high”. For the purpose of this public panel, it refers to the caution that needs to be exercised in using automated technologies in governance and designing safeguards to protect human rights.

**Pseudonymised data** – a technique that replaces or removes information in a data set that identifies an individual.

**Safe haven** - a secure place used to store particular research data, for access exclusively by approved colleagues. Strict safeguards control who can access medical and personal data for research. When researchers use this data, they must use IT systems with very high standards of security.

**Quantitative data** – data in the form of counts or numbers where each data set has a unique numerical value.

**Qualitative data** – information that cannot be counted, measured or easily expressed using numbers. It is collected from things like text, audio or images.

## **Session one: introduction**

### **The speakers**

- Scottish Government - **introduction to the Data & Intelligence Network**
- Nayha Sethi, University of Edinburgh - **introduction to data, data ethics and data justice**
- Stephen Peacock, Information Commissioners Office – **introduction to data protection**

### **Questions raised during session one:**

#### **The Data & Intelligence Network**

Questions (Q) and Responses (R). All responses provided by the DIN.

Q - If the DIN is trying to pull together a coherent platform, how do they separate trend data from identifiable data?

- R- The DIN does not provide a platform nor does it host data, so it has no need to separate trend data from identifiable data.

Q - Where does the DIN fit within the public sector – are they part of the Scottish Government? Are any private sector organisations involved in their work?

- R - The DIN is a collaboration made up of members across the Scottish Public and Not-for-Profit Sectors , including Health Boards/Agencies, Local Authorities, Academia and third sector. The DIN is led by dedicated team of Scottish Government staff who actively work with members of the network to help deliver projects, support communities of practice etc. Some services delivered by DIN members are supported by private organisations, and therefore there are occasions where the DIN engage with non-public sector bodies.

Q - The DIN work with quite a lot of different partners. What controls are in place to make sure data is processed/shared securely and appropriately (e.g. that data only used for specified purpose)?

- R - Projects go through an assessment process including an initiation document and, where appropriate, ethical workbook that helps identify and specify how data would be processed/shared securely and appropriately. Compliance with data protection requirements and other applicable legal frameworks as well as guidance from the Information Commissioner Office (e.g. Data Protection Impact Assessments, Data Sharing Agreements and Data Processing Agreements), is the

responsibility of the individual organisations involved in the project. The Scottish Government support team provides expertise and support to ensure the appropriate checks are in place.

Q - Could the DIN share some examples of the ethical considerations they have to make re. accessing data. And where is the proof that data is being used ethically?

- R - At the end of the project, the close out process will review whether the outputs including controls over the use of data have been successfully implemented.

Examples of ethical considerations the DIN members have taken in the past will be presented in workshops 2 and 3.

Q - At what level/who makes decisions about the ethical use of data?

- R - Decisions about the ethical use of data are taken at several levels. The first level is within member organisations themselves, when projects are proposed, or as issues or problems in a project emerge a first ethical assessment may take place. Should the DIN become involved we would work with members by going through the DIN ethical workbook to identify ethical concerns and determine how best to address these with the ultimate decision on how data should be used staying with the organisations sharing the data.

Going forward we envisage the public panel will help us by identifying overarching ethical principles reflecting the wider public's perspective so these are included in the decisions we make more about data use in Scotland, including the activities of the DIN.

Q - Are the public able to find out how DIN members are using their data and specifically what data are being accessed/used?

- R - Each member organisation will have different channels and platforms to publicise their data led projects. Additionally, when the SG support team agrees to support projects it will aim to be as transparent and open as possible about these projects through its newsletter, blogs etc.

Q - If there is a data breach what steps are in place to recover any leaked, personal data? And is it possible to fully recover it once it's 'out there'?

- R - The DIN doesn't hold any personal data. As indicated above, it is a responsibility of each individual member organisation to comply with data protection requirements and other applicable legal framework as well as guidance from the Information Commissioner Office.

Q - How are decisions made about what kind of data is shared with which kind of organisations? And does that create a possible conflict of interest between different members?

- R - When DIN members propose sharing data as part of a DIN project we will work through the proposals, highlight potential ethical considerations etc. We would also recommend research into whether a similar solution exists and whether the new project is the best solution and does not conflict with the aims and objectives of another member organisation. In general sharing between members is by agreement.

Q - How many people have access to our data, and how is this monitored?

- R - The DIN does not hold any data about data about individual (except for contact details of network members) and therefore does not monitor who accesses it. In projects that the SG Support Team helps deliver, the use of Data Privacy Impact Assessments (DPIA) and the DIN ethics workbook do ask project sponsors to indicate how many people would have access to data and how access is controlled/managed. Data access remains a responsibility of individual member organisations in compliance with data protection requirements and applicable legal frameworks as well as guidance from the Information Commissioner Office

Q - How can data be protected in transfer – especially (but not only) between private and public sectors? What role does encryption play?

- R - In every project we get involved in and support, we ensure data is held and transferred in a safe and secure way using recognised national and international standards.

Encryption is not always the best solution when transferring data. Encryption like other security tools needs to be used proportionately, and if used, should be as part of a well-structured data security solution. DIN members are best placed to agree their data sharing security needs with organisations they share data with.

Q - How does the government / public sector use AI in relation to data?

- R - The Scottish Government (SG) does not have any generic AI-specific internal policies and guidelines. However, we adhere to AI regulatory and policy frameworks already in place. Any public body in Scotland have to adhere to the Information Commissioner's Office (ICO) GDPR legal requirements of data.

They also need to take into consideration the ICO's Guidance on AI and data protection. SG also adopts the high level principles within the Scotland's AI Strategy and recommends any other public body to do so.

SG and public bodies in Scotland also collaborate with Centre for Data Ethics and Innovation (CDEI) to ensure that ethical considerations and the values that citizens want are reflected in governance and policy frameworks on the use of AI.

Q - Do different organisations (in the DIN) all follow the same safety standards – for example regarding internet security, or ethics?

- R - In every project we get involved in and support, we ensure data is held and transferred in a safe and secure way using recognised national and international standards. As regards ethics, there may be a first ethics assessment within member organisations themselves when projects are proposed, or as issues or problems in a project emerge. This initial assessment will be typically conducted in line with any ethics framework or guidance that organisations may have in place. Data ethics considerations are not widely applied yet.

Q - If some companies aren't always pseudonymising or anonymising data properly, how can government improve this situation?

- R - When we get involved and support a project we ensure that members organisation comply with Information Commissioner Officer guidance on anonymisation.

Q - If GDPR and DPA set policy, what is our role in this public panel?

- R - We envisage the public panel will help us by identifying overarching ethical principles that reflects the wider public's perspective. These will be included in the decisions we make more about data use in Scotland which are not covered by legislation and regulations, including the activities of the DIN.

## Data ethics

Q - If data sharing should be done 'to serve mankind', who decides what mankind is?

- R - University of Edinburgh response: Many justifications for using, collecting and sharing data hinge on the diverse benefits that data use can deliver. Concepts such as 'public interest', 'public benefit' and 'common good' play a key role in decisions about whether to authorise access to data not only in terms of ethical review of data use, but also, for example, with regards to setting aside legal requirements for seeking consent or anonymising data.

Each of these terms will be used in different ways depending on the context, but one thing they have in common is that they all appeal to notions of the common good, benefit, welfare or well-being of society. An important ethical question is indeed who gets to decide what constitutes public interest or public benefit. This might vary depending on a variety of factors including: the organisation holding the data, the type of data in question, the purpose for data sharing and who will be accessing it. Many organisations that hold health and care data will make assessments about public benefit as well as decision-makers across for example, University ethics committees, research ethics committees, public benefit and

privacy panels or data oversight groups. These committees or groups will be comprised of a variety of individuals with different expertise, skills and interests.

Further Resources: In their recent report [Data for Public Benefit](#) (2018) the Understanding Patient Data initiative held a series of workshops to understand what 'public benefit' meant in the context of data sharing involving personal data. More recently and in collaboration with the UK National Data Guardian, the initiative has released their report [Putting Good into Practice: A public dialogue on making public benefit assessments when using health and care data \(2021\)](#). The report is quite long but it is worthwhile reading the executive summary (particularly p 1 - 5).

- R - ICO response: I assume this question is who decides what serves mankind? This is a quote from a recital which is to be read alongside Article 1 of the GDPR. The recitals set out that the right to data protection is a fundamental right but not an absolute one, the right to data protection must be balanced against other fundamental rights. To comply with data protection law organisations must be able to demonstrate that their use of personal data is necessary and proportionate and does not result in a high risk to rights and freedoms. When deciding whether to share data organisations must weigh up the risks of not sharing against the risks of sharing. The ICO as the UK regulator of data protection can take regulatory action where data protection law has not been complied with the courts the final arbiter.

## Data protection

Questions (Q) and Responses (R). All responses provided by the ICO.

Q - If data is shared for research purposes: Is it de-identified? And have people consented to their data being shared for this purpose?

- R - The short answer is it depends!

Is it de-identified?

Data protection law says that: You should only process personal data that adequate relevant and limited to what is necessary for your purpose.

In some cases it will be possible to conduct the research using entirely anonymous data (in other words data that is no longer personal data because key identifiers have been removed so that individuals are no longer identifiable or the chances of identifying any individual is sufficiently remote).

For other research it may be necessary to retain some identifiers so, for example, different data sets can be combined (linked) e.g. information about your health that might be held by the NHS with information about your education which may be held elsewhere)

or if you are tracking an individual's progress over time. Where this is the case the researchers can add a layer of protection called pseudonymisation (or de-identification). This would be where an identifier like someone's name is replaced with a reference number. Pseudonymisation means that people are not identifiable from the dataset itself. However, they are still identifiable by referring to other, separately held information. This gives individuals a layer of privacy and reduces risk of harm whilst not obstructing the research.

ICO guidance says that where personal data is being used for research purposes researchers should: ensure that you do anonymisation or pseudonymisation at the earliest possible opportunity, ideally prior to using the data for research purposes.

Have people consented to their data being shared for this purpose?

This depends upon the research being conducted. Data protection law recognises the importance of scientific and historical research to society and through a set of provisions called the research provisions. These allow data collected for one purpose to be shared for research purposes without the need to collect fresh consent or a specific legal requiring or allowing the sharing provided certain safeguards are in place. This includes suitable pseudonymisation.

Q - Will there be changes/deregulation of data protection rules in the UK as a result of Brexit?

- R- The Data Protection and Digital Information Bill was introduced to Parliament on 18 July 2022. You can read the Bill [here](#). If this becomes law it will make some changes to the current data protection law. The proposals are for a more flexible, outcomes-focused regime that supports responsible data use and innovation while still protecting individuals' rights.

Yesterday (03 October 2022) at the Conservative Party Conference, Michelle Donelan, Secretary of State for Digital, Culture, Media and Sport announced her approach to data protection reform which includes introducing a new UK data protection framework. It is not clear at present what changes will be made to the existing proposals.

Q - How do you protect people's privacy when people are working from home? Is there greater risk of breaches as a result of this?

- R- Working at home may present new or different risks from office based. Employers should have measures and policies in place however to ensure that these risks are managed and reduced and that sufficient protections are in place and data protection law is complied with. At the beginning of the pandemic we published some guidance for employers on this: [Working from home | ICO](#)

Q- What is the ICO's audit process?

- R- ICO audits assess whether an organisation is following good data protection practice and meeting its data protection obligations. Following an audit, a report will be produced that sets out recommendations for improvement. The ICO takes a risk based approach to identifying which organisations it audits. We focus on those areas we feel we will have the biggest impact and organisations who would benefit the most from an independent assessment of their compliance with data protection legislation. More information on the process can be found here: [A guide to ICO audits](#)

Q - How is special category data used differently and why is it therefore defined as special category?

- R - Special category data includes:
  - personal data revealing **racial or ethnic origin**;
  - personal data revealing **political opinions**;
  - personal data revealing **religious or philosophical beliefs**;
  - personal data revealing **trade union membership**;
  - **genetic data**;
  - **biometric data** (where used for identification purposes);
  - data concerning **health**;
  - data concerning a person's **sex life**; and
  - data concerning a person's **sexual orientation**.

It is not that special category data is used differently that means it is defined as special category. Rather it is because data protection law gives it greater protections. This is because use of this data could create significant risks to the individual's fundamental rights and freedoms. For example, the various categories are closely linked with:

- freedom of thought, conscience and religion;
- freedom of expression;
- freedom of assembly and association;
- the right to bodily integrity;
- the right to respect for private and family life; or
- freedom from discrimination.

## **The panel process**

Q - Can we see the presentations in advance to help us get our head around the material before we meet?

- R (Ipsos) - thank you for this suggestion and it's a great idea. We will look at doing this for future sessions.

## **Session Two: Past Projects (pt 1)**

### **The speakers**

- Scottish Government – **the shielding list project**
- Scottish Government – **the CURL project**

### **Questions raised during session two:**

#### **The shielding list project**

Questions (Q) and Responses (R). All responses provided by Scottish Government.

Q - What's happening with shielding data now? Is it being kept? How long for? Is it being updated and how?

- R - Use of the shielding list (also known as highest risk list) ended on 31st May 2022. It will be kept in line with NSS retention policies but will not be amended or updated. Whilst the data on individuals who were on shielding lists is not being updated the underlying data on health conditions of individuals will be updated on GP systems.

Q - What other third parties (apart from employers) received the shielding data?

- R - We are not aware of any other third parties receiving or accessing the shielding data.

Q - What happens if data is missing for an individual?

- R - When it was operational It would depend on many factors: What type of information was missing? When was the missing data identified? Who needed to be contacted to get the correct information etc.? However as the shielding list is no longer being updated any missing data will not be changed as it provides a record that data was missing.

#### **The CURL project**

Questions (Q) and Responses (R). All responses provided by Scottish Government.

Q - Were any private sector organisations involved in handling the data?

- R - No. There are no private companies or organisations involved in handling or processing the CURL data.

Q - Why wasn't electoral roll data used to update the addresses?

- R - The underlying law (Representation of the People (Scotland) Regulations 2001) that requires the Electoral Register (roll) to be created doesn't allow the register to be used for anything other than organising and running elections and a very limited range of other Local Authority activities.

Q - What is the main purpose of CURL? Are there any particular problems you feel CURL would be a good solution for?

- R - CURL was developed to understand where discharged hospital patients were transferred to and to understand COVID-19 testing in Care Homes. CURL has also been used to understand at the geographic spread of outbreaks and vaccine take-up across Scotland. CURL could be used to, more generally, look at the spread of different types of health conditions and how they relate to environmental factors, housing conditions and other deprivation indicators.

### **Wider ethics and the role of Network**

Questions (Q) and Responses (R). All responses provided by Scottish Government.

Q - Who has the ultimate say on how data is handled and used, particularly with regards to ethical decisions?

- R - The Data Owner and the Data Access Authority control who can access personal data, how it is accessed and specify the security features the computer environment that needs to have. Ethical checks should be carried out by the organisation the person wanting to access the data works in. For example, university researchers follow their internal ethics approval process, while Scottish Government statisticians use the UK Statistics Authority process. The DIN has an Ethics Workbook that all the Network members requesting our help should fill in. We would like to ask you, the citizens, who do you think should review the Ethics Workbook.

Q - Will I be told if my data is involved in a data breach?

- R - When the risk is considered high risk, all individuals must be told. The Data Protection Act 2028 requires this.

Q - Do public sector bodies get fined like private sector ones do?

- R - Yes

Q - Is there ever compensation for people whose data has been compromised?

- R - That would be a matter for a court or maybe the Information Commissioner's Office (ICO).

Q - Who do you turn to if data is mishandled?

- R - The organisation Data Protection Officer (DPO), the organisations Senior Information Risk Officer (SIRO) and ultimately the ICO.

Q - What are the actual benefits for society of analysing anonymised data?

- R - This is a very big question! Briefly, by letting Public Sector and university researchers use anonymised, real data about individuals, they can get a true understanding of how people go about their lives and how this could be affected by different policy decisions. We can also use research findings to help colleagues design better policies. Because the data is often routine information that public sector organisations already collect, it is less intrusive and costly than collecting the same information by other means, such as through large-scale surveys. It allows entire populations, or specific parts of the population to be studied, reducing common problems with gaps in data often encountered in surveys.

Q - What's the step by step process for ethical approval of a project? How do you decide on a project when there's no obvious right or wrong position?

- R - At the moment, all Network members requesting our help should fill in the Ethics Workbook which explores the risks and benefits of the project at different levels. We do not yet have a formal process for assessing the completed Ethics Workbooks, and would like to ask you what you think this process should look like from the citizens' perspective.

Q - As we review all these projects, what should we be worried about? What are other people worrying about when it comes to data being used in these ways?

- R - In reviewing projects we would suggest, amongst other factors, that panel members think about whether the benefits of a project to individuals or the wider community outweigh any risks. As the nature and scale of the projects vary so much there will be a range of worries and concerns. By bringing their own experiences and concerns Panel members will help projects see the "bigger picture" and reflect on wider concerns.

### **Session Three: Past Projects (pt 2)**

#### **The speakers**

- Duncan Buchanan, Research Data Scotland – **the equalities and protected characteristics project**
- Scottish Government – **the Ukrainian Displaced People project**
- Anuj Puri – **an outside perspective: reflections on the ethical issues**

#### **Questions raised during session three:**

## Equalities project

Questions (Q) and Responses (R). All responses from Research Data Scotland.

Q - How do the Scottish Government keep the data collected (i.e. from the Census) up to date with more recent data? And how do they ensure they do this correctly?

- R - The census is collected and stored separately as a one off exercise every 10 years. To use it along with other more up to date data sources, like from the NHS, requires the quite complex data linkage work I described in my presentation. Once linked together, you can find the most up to date records for individuals. However to keep this exercise up to date requires a plan to continually 'refresh' this data linkage at regular intervals, say every 3 months or every year. We haven't got that plan yet having just completed the initial data linkage work. But it's something we need to consider once we can demonstrate the value, and security, of having the up to date equalities all together in one place. (Response provided by Research Data Scotland)

Q - In what ways will the data collected be used now, and how is that decided?

- R - The first step is to run a so called 'proof of concept' project where the dataset is used by one public sector organisation to analyse equalities data for real. We have a couple of organisations interested but are still arrange something that can be done fairly soon and is of pressing need. This may run smooth or may highlight further improvements needed. Further to this we need to consult with public bodies and researchers on whether it meets their needs for equalities duties and monitoring. If successful we need to return to the independent scrutiny panels (one for NHS data, one for government data) with a plan for how, and to whom, the dataset will made available in future. However, the basic model will follow the model currently in operation for access to sensitive public sector data for research. That model involves each project applying for access and being assessed based on the 5 safe's framework used across UK: safe people, safe projects, safe settings, safe outputs, safe data. This normally is done by the independent scrutiny panels. (Response provided by Research Data Scotland)

Q - To what extent does anonymity, or pseudonymisation, of data compromise data quality?

- R - It does not affect data quality of the characteristics (like age, ethnicity or religion) directly. The data quality issues can arise from the process of anonymisation which involves attempting to identify the same individual across different datasets and systems, e.g. census, or school, or NHS. Because this relies on personal information like names, addresses, date of birth etc, any variation in how these are recorded across systems (e.g. mistakes, spelling differences) increases risk you can't identify the same person appearing in different systems. So individuals could get missed or are assumed to be 2 different people. So when you anonymise them and remove personal information there's no way of knowing they could be the same people or that some people are not appearing. There is always a percentage in this bracket given the numbers involved but it's usually pretty low. Lots of in-depth academic work has been done on developing the

algorithms used in data linkage methods like this but it's definitely something that people using the data need to be aware of when doing their analysis on the anonymised data.  
(Response provided by Research Data Scotland)

## Ukrainian Displaced People project

Questions (Q) and Responses (R). All responses from Scottish Government.

Q - What differences are there between Scotland and other UK nations re. how approach data on Ukrainian refugees and hosts?

- R - Immigration is not devolved, so all visa applications are processed by the Home Office following UK Gov policies. A UK wide system exists but Scotland and Wales decided not to use it as it would mean being closely coupled with all UK policy. Scotland and Wales both implemented a 'super sponsor' scheme where the government is the visa sponsor and is responsible for matching the displaced person(s) with temporary accommodation. This has been the main policy difference. The approaches to data is similar across all 4 nations as local authorities are responsible for housing and child services throughout the UK. A difference in Scotland was the decision to use the NHS NSS call centre for initial contact with people applying under the super sponsor scheme. This was already in place for dealing with COVID and was reused to enable Scotland to contact large numbers of people with immediate effect.

Q - Who has access to data on hosts/refugees? Which delivery agencies/parts of the council?

- R - NHS NSS Call Centre get visa data for super sponsor scheme only.

Local Authority Housing Departments get visa data for private sponsor scheme only and host data to carry out property checks and background checks on hosts.

Local Authority Child Services get visa data for super sponsor and private sponsor schemes long with host data in order to carry out safeguarding checks where minors are involved.

COSLA, Local Authority Housing Department and Scottish Government Matching Team also receive augmented data from the previous steps with additional data that helps them match displaced people with hosts.

There is some additional data concerning finance, education, etc but this is collated and aggregated data with no personal data.

Q - How is data security managed on the Ukraine project?

- R - We have strong processes in place for authentication and authorisation (including Multi Factor authentication). All requests for new data or changes to existing data are subjected to governance processes. We are currently using a standard Scottish Government system for sharing data which is already compliant with cyber and GDPR

standards. As we develop new systems to speed up the matching process and reduce the time people spend in temporary accommodation, these are subjected to the same assessments to ensure the ongoing security of the data.

Q - Is there any risk to 'group privacy' / a risk to the Ukrainian community from how data could be used?

- R - This was not initially considered as the main priority was providing refuge and safety to very vulnerable people. It is something we need to consider again as we start to plan longer term integration.

Q - What level of anonymity (if any) was there in this project?

- R - When it comes to processing visa application, hosts and matching, none of the data is anonymised as we need to process each individual case with the personal data. All other data, including published statistics are aggregated to ensure anonymity.

## Reflections on the ethical issues

Questions (Q) and Responses (R).

Q - Can we have definitions of Constitutional privacy, Group privacy, Information privacy, Algorithm governance, Contextual integrity, Contestation, Data minimisation, Informed consent, Precautionary principle?

- R - **Ipsos**: with help from Anuj, we have added all these definitions to the glossary (see pages 5-7).

Q - How does Anuj think the precautionary principle could be embedded in ALL data sharing projects?

- R - **Anuj**: When it comes to incorporating safeguards in data sharing projects, precautionary principle is not the only way forward. Other regulatory approaches may include a risk based approach where the project organiser(s) has to identify the risk and harm based approach.

None of these approaches may by themselves effectively balance the interests of the data subjects with the benefits of data sharing. Hence, recent research encourages the adoption of a nuanced approach that aims to align the uses of data with the needs and rights of the communities reflected in it. While selecting the regulatory framework, we must be guided by the nature of the data involved and its impact on the rights of data subjects. When it comes to data sharing, not all harms suffered on account of violation of privacy can be monetarily compensated, hence depending upon the sensitivity of data involved we need to proceed cautiously both on account of confidentiality involved as well as the inferences that can be drawn on the basis of such data. In order to embed effective safeguards in a data sharing project, it would be helpful to carry out a privacy

impact assessment of the project and assess the nature of data involved, risks to privacy and other human rights.

Q - What does data minimisation mean in practice?

- R - In practice from an organisational perspective, data minimisation would require setting out of clear goals for which data is required, assessment of minimum data required to achieve those goals, development of protocols that restrict access to data and sharing of data, and placing time limit on storage of data. You may find this guide from the UK ICO helpful:

<https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/data-minimisation/>

## Broader issues

Questions (Q) and Responses (R).

Q - How do you at Ipsos handle our data as panellists? What information do you hold on us and where do you get it from? How were we chosen?

- R - Ipsos: great question! You were randomly selected to join this panel and this involved a few stages which we describe below. You might remember receiving an invite in the post from the Sortition Foundation – they specialise in this type of recruitment and you can read more about them on their [website](#). They sent invitations to a random selection of addresses across Scotland – those addresses were taken from the Postal Address Finder (PAF) which is owned by Royal Mail. Organisations who want to use the PAF have to request and purchase it from Royal Mail. Your address was randomly selected to receive an invitation.

When you signed up for the panel you will have completed a form which asked questions about you (like your age, gender, ethnicity etc). This information was used by the Sortition Foundation to make sure that our final selected panel is broadly representative of the Scottish population. With your permission, this information along with your contact information, was passed along to us at Ipsos once you were confirmed as a panel member.

We have used this information to keep in touch with you about the panel. We have also collected further information from you (such as your bank details) to help us run the panel. As we record the sessions to make sure we capture everything, this means further personal information is collected (i.e. your voice is considered personally identifiable information).

All your personal data is held securely on our systems and is only accessible to a few team members. It is never passed to anyone outside of the Ipsos research team and is securely deleted from our systems once the project is complete. You can read more

about how we use your personal data in the privacy notice and information sheet that you received in your welcome pack. We have published these on the online community as well so you can have a read any time. If you would like to receive another copy by email, please just get in touch!

## **Session four: principle forming**

### **The speakers**

None

### **Questions raised during session four**

Questions (Q) and Responses (R). All responses provided by the Scottish Government.

Q - Can you request what data the Scottish Government holds on you?

- R - Under the UK Data Protection Act, each citizen has the right to see personal information organisations hold about them, request correction, deletion or restrictions on the use of their personal data: <https://ico.org.uk/your-data-matters/your-right-to-get-copies-of-your-data/>

Q - Are private sector organisations subject to all the same data rules as public sector ones?

- R - Yes they are.

Q - Is the data ever truly anonymous when being analysed by government? (thinking about the drugs and alcohol case study example which used health data, police data etc)

- R - Given enough time, it may be possible to re-identify individual people from an anonymised dataset. This is strictly prohibited by the civil service code of conduct and professional standards. The type of anonymised data and type of analysis applied on different projects would impact the probability of a person or a group of persons potentially becoming identifiable. This guide from the ICO provides lots of good and trustworthy information on anonymisation [anonymisation-intro-and-first-chapter.pdf \(ico.org.uk\)](#).

# 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 market research specific standard that supersedes BS 7911/MRQSA and incorporates IQCS (Interviewer Quality Control Scheme). It covers the five stages of a Market Research project. Ipsos was the first company in the world to gain this accreditation.



## Market Research Society (MRS) Company Partnership

By being an MRS Company Partner, Ipsos endorses and supports the core MRS brand values of professionalism, research excellence and business effectiveness, and commits to comply with the MRS Code of Conduct throughout the organisation. We were the first company to sign up to the requirements and self-regulation of the MRS Code. More than 350 companies have followed our lead.



## ISO 9001

This is the 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

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



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

Ipsos is required to comply with the UK GDPR and the UK DPA. It covers the processing of personal data and the protection of privacy.



## HMG Cyber Essentials

This is a government-backed scheme and a key deliverable of the UK's National Cyber Security Programme. Ipsos was assessment-validated for Cyber Essentials certification in 2016. 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.



## Fair Data

Ipsos is signed up as a "Fair Data" company, agreeing to adhere to 10 core principles. The principles support and complement other standards such as ISOs, and the requirements of Data Protection legislation.



© Crown copyright 2024

**OGL**

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit [nationalarchives.gov.uk/doc/open-government-licence/version/3](https://nationalarchives.gov.uk/doc/open-government-licence/version/3) or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk).

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

This publication is available at [www.gov.scot](http://www.gov.scot)

Any enquiries regarding this publication should be sent to us at

The Scottish Government  
St Andrew's House  
Edinburgh  
EH1 3DG

ISBN: 978-1-83601-043-2

Published by The Scottish Government, August 2024

Produced for The Scottish Government by APS Group Scotland, 21 Tennant Street, Edinburgh EH6 5NA  
PPDAS1430446 (08/24)

W W W . g o v . s c o t