February 2023

Creating Resilient and Trusted Data Systems
Executive Summary

A public dialogue on effective data environments for emergency and non-emergency situations

Conducted on behalf of the Royal Society

Contents

Executive Summary

Background

The Royal Society commissioned the deliberative engagement specialists Hopkins Van Mil (HVM) to conduct a public dialogue on creating resilient and trusted data systems. The full process including design and fieldwork ran from November 2021 to February 2022. It was commissioned as a mainly face-to-face dialogue, but as Omicron arrived in the UK, it moved to an online deliberative process.

The ability to access quality data by scientists and government decision-makers rapidly has been shown to be essential during the Covid-19 pandemic. In the early part of the pandemic there were challenges in accessing and sharing quality data in a timely manner which hampered the ability of advisers and decision-makers to understand the situation and recommend actions.

Huge progress was made to respond to this challenge, but some critical questions remain which include:

- Can the systems we have created now help us in a future pandemic?
- Have the systems been established in ways that enable them to be used in a trusted way outside of emergencies?
- Are we any better placed to have a data-led response to other emergencies?

The Royal Society consider it critical to understand the public's views in an exploration of these questions and position public consideration at the heart of policymaking. This dialogue was therefore commissioned to consider how to build a future data-led response to emergency and non-emergency situations faced by society.

Dialogue aim and objectives

The dialogue set out to inform the work of the Royal Society's Resilient Data Systems for Emergencies programme. This aims to identify how to build an amenable data environment for the UK where quality data sits alongside robust mechanisms for enabling access to it. Such a system would be suitable for both emergency and non-emergency situations.

The research question explored by participants is: how do we develop a system for using data which is resilient, effective and trusted in emergency and non-emergency situations? To reflect on this dialogue participants were supported to:

- Explore levels of awareness of data systems, including understanding of the current data landscape, data flows, data use and data governance in different emergency situations, and during non-emergency situations
- Define what emergency situations are, and the different types of emergencies

- Explore views, expectations, and concerns around data use, flows, and data governance during contrasting scenarios of different emergency situations e.g., health emergencies, environmental emergencies (both short-term events such as flooding versus longer-term climate emergency response), and non-emergency situations, setting out where the main ‘trade-offs’ and ‘win-wins’ may be
- Explore how data systems can exacerbate inequalities and how future systems can be made more inclusive
- Uncover how views may change within different situations, with regards to the use and access of different types of data, through different organisations, with regards to an absence or shortage of data, how trade-offs may change and where new priorities emerge
- Create recommendations which highlight where there appears to be unanimous and clear priorities for action as well as pulling out the nuance of context-specific recommendations and conflicting points of view.

Participants drew on stimulus with a global perspective and did mention the international dimensions of data systems. However, this dialogue reflects the views of participants from the UK mainly reflecting on UK data systems. International comparisons are therefore limited.

Dialogue stimulus materials

Before taking part in the dialogue participants were sent a workbook[1]. This was presented in two parts. The first gave joining instructions for taking part in a public dialogue, including on using zoom and guidance on joining the online homework space. The second provided content materials including a data systems jargon buster, programmes for each workshop and emergency and non-emergency scenarios and other stimulus materials.

As participants joined the online homework space they could access electronic versions of the workbook and additional contextual material. This included an overview of how health data systems operate, including information on, for example: health and care records; the use of confidential information in health care; and the kinds of data used for test and trace mechanisms during the pandemic.

This dialogue was commissioned during the global Covid-19 emergency. The experiences from the pandemic shaped some of the stimulus materials presented to participants and was a focus for some sessions. Participants equally considered other emergency situations such as:

- Local flooding incidents
- The climate emergency, particularly the impact of significant and repeated heat waves
- Public health emergencies such as an ageing population living in inappropriate housing, particularly for those with multiple morbidities and who are experiencing a bigger care and poverty gap because of the pandemic.

---

[1] Appendix 1

All participants received the same information during the dialogue, with slight adjustments for location. For example, climate change projections for 2030 were given to participants based on their own location[2].

Recruiting participants

The Royal Society commissioned a public dialogue[3], a deliberative and qualitative research method which works with smaller samples of people than are found in large scale quantitative surveys. The method is selected because commissioners wish to gain a depth of understanding on participants' attitudes, views, beliefs, values and needs which is not possible with those methods which involve more participants but which are not deliberative or qualitative. Programmes similar to public dialogue include Citizens' Juries and Assemblies.

Participants are recruited through a process of purposive sampling, as distinct from random sampling often used in quantitative research, to involve a selection of people who have the potential to reflect a wide set of views, values and demographies. Participants are not self-selecting but join the dialogue based on demographic information agreed by the project team and set out in the recruitment specification[4]. This includes purposefully sampling for those from rural and urban locations; for a range of ages and life stages. We sample for a balance of genders. A boosted sample was used to ensure that people disproportionally affected by emergency situations, including disabled people, those from minoritised ethnic groups and from lower socio-economic groups were over-represented in the sample. The sample was produced using relevant Office for National Statistics, local authority and 2011 Census data to broadly reflect the locations from which the participants came.

Recruitment was carried out[5] using on-street methods, through community groups and (as a back-up) from agency panels. We exclude those who have taken part in public dialogue, Citizens' Juries or Assemblies in the last twelve months to avoid research fatigue and an over-familiarity with the process. Participants were recruited to take part in one of five public dialogue groups from a thirty-mile radius of Leeds, Glasgow, Cardiff and Belfast. In addition, one group was recruited from across the UK. Twenty to twenty two participants were recruited to each group, with 111 participants taking part in total. To ensure we involved people with a range of perspectives on data we asked participants in the recruitment process two questions about their views on data sharing and on social media usage.

The public dialogue method

Dialogue works when participants interact on a level playing field with specialists. This specialist evidence is then viewed through the lens of participants' own lived experience, acting as a provocation which leads to rich and powerful insights.

---

[2] Using, for example, What will climate change look like near me? BBC/ Met Office, July 2021

[3] More detailed definitions of public dialogue are available from Sciencewise

[4] The Recruitment Specification used for this project is available at Appendix 2

[5] We work with the specialist agency Roots Research to recruit participants

This process leads to an in-depth understanding of what people value, what they are concerned about, their priorities and the principles they apply to this prioritisation. HVM facilitators are key to gaining this understanding. They ensure there is a balance in small group discussions which allows people freedom to express their views whilst not allowing the process to lose the important focus on the dialogue scope or for the exercise to be derailed. This report sets out the findings that have emerged from this public dialogue process. Recruiting a diverse group of people to the dialogue ensures we hear, and participants respond to, a diversity of views. Dialogue participants learn from the process. They are influenced by the speakers and by their fellow participants. For many participants the dialogue was the first time they had thought to any degree about the data systems used to inform public and private policy.

The fieldwork took place between December 2021 and February 2022. Dialogue participants heard from expert speakers who gave contextual material on data systems, including on NHS data systems, UK statistical regulation and data for global emergency and risk planning. Each location had one live speaker. Presentations were recorded at workshops and shared with participants from all locations in the online homework space. This allowed all participants to review all the presented material. The dialogue process for each location is set out in Figure A.

Figure A: Public dialogue outline method

In was HVM's intention to deliver four of these five sets as face-to-face dialogues, with the UK wide cohort running using online methods. Unfortunately as a result of the onset of the Omicron variant in December 2021, and in line with public health guidance at that time, the decision was made to re-purpose the workshop design for online delivery. We retained the key essentials of a face-to-face dialogue so that each group took part in two rounds of workshops. For some locations this comprised two week-end workshops, two to three weeks apart, for others it comprised a combination of evening and week-end workshops, and others only evening workshops. All participants spent twelve hours in workshop discussions with an additional two hours using an online space in their own time.

The dialogue process included the use of the following tools:

- Mentimeter, an online polling tool, used in the workshops to gain a snapshot of views
- Jam Boards for facilitators to take visible notes during the workshops which participants can amend and build on as their discussions develop
- Recollective: an online qualitative research tool, which enables participants to review and comments on materials, answer questions such as their views on data sharing, and to reflect on their lived experience in their own time outside of workshop discussions.

About this report

Public dialogue findings cannot be taken to be statistically representative of the general population. However, they do uncover participants' views and the values, beliefs, experiences, interests and the needs that underlie them. As such we refer throughout this report to the views of dialogue participants rather than making any broader claims of being able to extrapolate the findings to the UK population.

The online dialogue workshops generated sixty hours of audio recordings. These were transcribed and, with the materials from Mentimeter and Recollective were analysed using NVivo software. Our reporting includes summaries of the analytical work participants did during the process combined with researcher analysis resulting from a comprehensive review of the dialogue data. We make the difference clear throughout the report.

HVM applies grounded theory to our analysis of public dialogue deliberations. We build theories from what we have heard rather than having a preconceived hypothesis to test. We make use of Sciencewise Guidelines for Reporting (July 2019) and the evaluation of previous public dialogues to inform our work. Throughout the process the HVM coding, analysis and writing team have maintained a rigorous approach and held frequent sense-checking sessions to mitigate against researcher bias.

We use terms such as 'a few', 'many', 'several' or 'some' to reflect areas of agreement and difference. These should be considered indicative rather than exact.

It is important in any dialogue process that the report reflects the voices of participants. Therefore, we have used short quotations from those who took part in the dialogue, drawn from the transcripts, to illustrate the analytical points being made and to emphasise main points. We have also used longer 'lived experience' quotations throughout the report which describe in participants' own words an experience which highlights a relevant data system experience. Some quotes have been edited to remove repeat or filler words. There have been no other edits which might distort the meaning intended by participants. In conducting the analysis and reporting on the findings HVM researchers have made judgements about which quotations to include. These judgements are based on a respect for what participants shared and the seriousness with which they took their role in the dialogue. Quotation selection was also made in relation to what best reflects the key themes raised, including a diversity of voices, and highlighting the key points from a participant and researcher led analysis.

The dialogue findings

This dialogue has revealed what matters to participants when they consider resilient, effective and trusted data systems. Findings are divided into two sections: section A sets out the key findings by main theme; section B builds on this and reveals further key findings explored via the 'data conundrums' participants' identified alongside the potential solutions to these dilemmas. This section also highlights areas which are valuable for further exploration between specialists in the field and publics. This is followed by recommendations made by participants in the dialogue.

Findings by theme

Theme 1: Assumptions, surprises and early reactions

Participants on joining the dialogue assume that:

- Data systems refer to privately owned and commercially operated systems
- Data collected for commercial purposes is sold on without thought, regulation or any process of redress for harms such as scams, hacks and data related fraud
- Data systems rely on 'me' putting 'my' data into the system and therefore barriers to doing this create inequalities in data systems
- Data is easily accessible and widely available to those who might want, need or wish to use it for public benefit, commercial purpose or for criminal activity
- When public sector data systems do come to mind it is assumed at first that these are linked and inter-operable e.g. that hospital consultants will be able to see GP records.

Finding 1: Participants are surprised when they consider how much data is shared unthinkingly on a daily basis. This leads to astonishment that people, including themselves, are so trusting of organisations collecting data, particularly when it is not clear how the data will be used.

Participants considered where they sit on the scale we described as 'Keep it close Kieron: data about me is private and shouldn't be used be used beyond its original purpose - even in emergency situations' at one end and 'Give it away Greta: data about me should be used or planning for and improving services – whatever the situation' at the other. Most people situate themselves in the centre of this scale (figure B). They feel:

- There is a balance to be struck between data privacy and data availability
- Data should be used 'correctly' and 'appropriately'
- Acceptability comes from being clear about who will have data access for what purpose(s)
- People are less willing to engage with data systems if they feel there is the potential for exploitation of citizens; surveillance of society by government; or to justify actions they believe to be contrary to the public good.

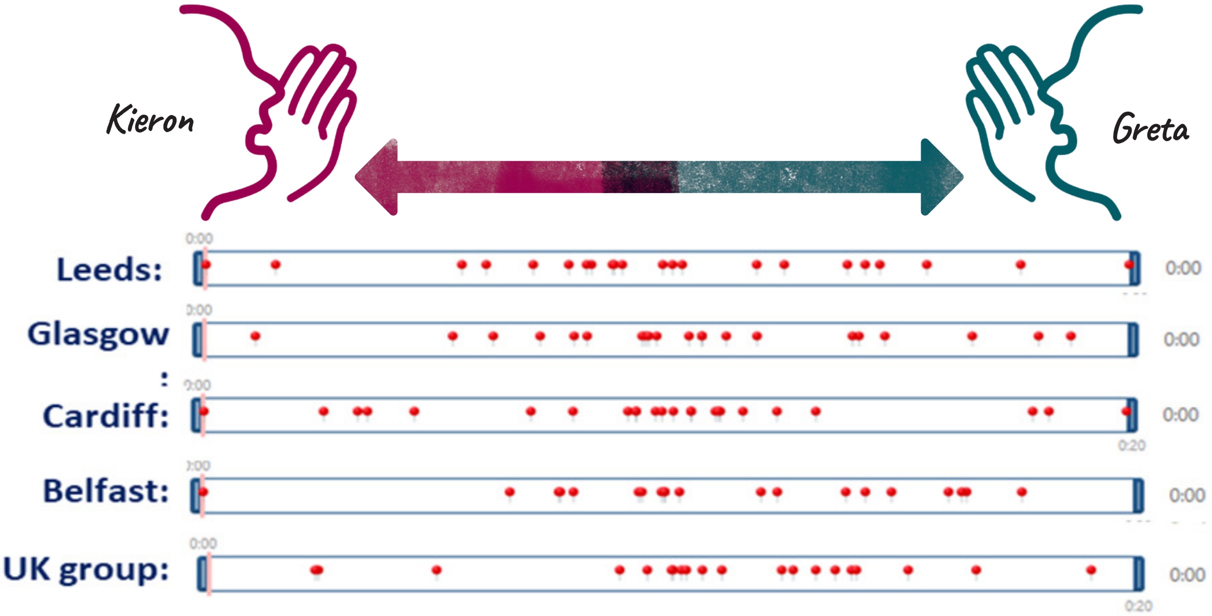


Figure B: Reactions to data use

Finding 2: Participants who put themselves towards Kieron say they are concerned that they have no control over how their data is used; feel exposed to harm; and do not trust those who manage data systems not to exercise power and control over those whose data is collected. Participants who lean towards Greta believe that data should be used for public good, particularly to improve public services and emergency responses.

Finding 3: Participants are concerned that the parameters of an emergency are clearly defined. Figure C presents the words participants use to define an emergency. They believe society needs to understand the value of data systems in an emergency in order to assess what is appropriate data use in non-emergency situations.

Figure C: Participant emergency definitions

Theme 2: Trust and transparency

Finding 4: There is a strong sense expressed by participants that data should be collected and used for defined purposes – even if those purposes are not entirely evident e.g. protecting society from future emergencies. Mistrust of data systems arises when this clarity of purpose does not appear to exist.

Participants express high levels of trust in frontline professionals such as health, care and environmental protection professionals as well as public/ academic experts in data. Lower levels of trust are expressed in those with an ‘agenda’ for data which could conflict with public good. This includes commercial and party political objectives (figure D).

Figure D: The spectrum of trust

Finding 5: Mistrust of data systems and those who manage them is characterised by exploitation and misuse. Trust is characterised by expertise and public good.

Finding 6: Participants identify seven facets of trust (Figure E), essential elements which must be woven in to data systems to achieve public credibility and durability.

- Free of exploitation: public benefit first and foremost
- A diversity of organisations & experience involved
- Clear & tangible benefits are clear and well communicated
- Data used by those who need it for a specific purpose
- Monitoring data usage is part of effective governance
- Use data to model a range of scenarios: in an emergency showing options build trust
- A recognition that building trust takes time

Figure E: The seven facets of trust

Finding 7: Transparency depends on having a clear understanding of what the data collected actually achieves. Participants feel this is largely hidden from public view and this lack of awareness of the purpose of data systems leads, they believe, to missed opportunities to reassure society of the value of data and build trust in data systems.

Theme 3: Balancing inclusive data systems with those that protect individual privacy

Finding 8: Making data systems inclusive is a key priority for participants. They articulate a number of impacts that data systems not designed with inclusion and diversity in mind can have on people's lives. These are set out in figure F.

Not being able to access services in non-emergency situations e.g. healthcare

Not being part of the system which would help them in an emergency situation – missing from data sets

People becoming victims of online crime due to a lack of understanding of the risks of online harm

Discrimination in data sets e.g in relation to employment

Figure F: Impacts of data systems that are not inclusive

Finding 9: Participants agreed that, in an ideal world, data systems would demonstrate the hallmarks of inclusivity. Including being accessible, representative and giving a fair interpretation of the data (figure G)

Hallmarks of inclusive data systems

- ✓ Accessible to all
- ✓ Representative of everyone
- ✓ Prioritise the vulnerable
- ✓ Non-discriminatory
- ✓ Design with inclusion and accessibility front of mind
- ✓ User-friendly and simple
- ✓ Available in different formats
- ✓ Available in different languages
- ✓ Engagement and support
- ✓ Fair interpretation of data

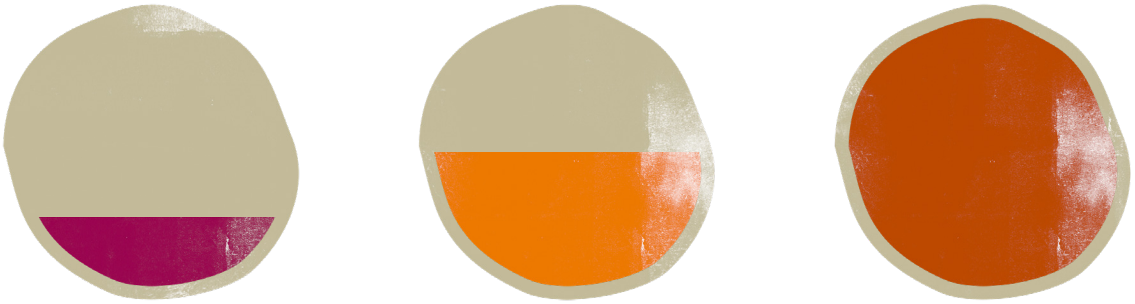
Figure G: Ideal hallmarks of inclusivity

Finding 10: Participants express concern about data privacy, particularly in industry-led data systems, but across all systems. They believe that:

- Personal data should be private unless de-identified
- Opt-in/ opt out choices are important
- Individuals should have the right to know what data is included in the system and to remove data from it
- Data privacy must still matter in emergency situations, even though use of data might be more urgent and more personal and identifiable data might be needed to address the emergency.

Theme 4: the need for data in emergency situations

In discussing what is needed from data systems in emergency situations participants tend to fall into three groups (figure H).



Cautious

The data needed in an emergency should already exist, no need for extended data access powers in an emergency.

Willing

Data should be accessible in an emergency - as long as society is aware of the benefits and data protection is robust.

Ambitious

Anything and everything should be done, including unrestricted access to personal data, if this is in the public interest and will end the emergency situation.

Figure H: Three main attitudes towards data systems in an emergency

Finding 11: Emergency powers to access personal, identifiable and sensitive data should not extend beyond the emergency situation and become 'normalised'

Finding 12: Participants conclude that connected data systems are of value to society in both emergency and non-emergency situations. They call for more co-ordination and inter-operability between systems, particularly those which deliver public benefit such as health, care, housing and education.

Participants recommend that:

- A shift is needed to recognise that linking data systems and fostering a spirit of collaboration between those who manage them is likely to produce greater public benefit in their use
- This shift will also minimise the burden on society in collecting and recollecting data for different purposes and bring specific public benefits in key social and economic areas such as health, social care, education and housing
- Public reassurances need to be made on the purposes for which data is collected and shared

Theme 5: Resilient data systems

Finding 13: Participants see learning throughout the system as an important aspect of resilience. They describe a very simple cycle of learning (figure J) which reflects their desire for an effective, resilient and trusted data system to learn constantly. This reflects a need for data systems to be open to what's not working well and identifying areas for improvement.

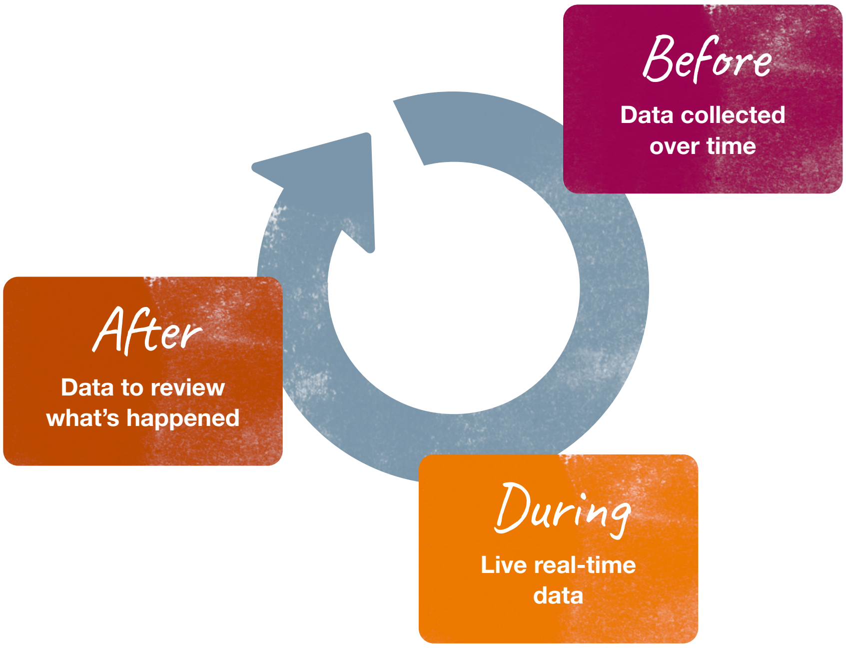


Figure J: The data system cycle of learning

Finding 14: Four specific learning tools are identified by participants which they believe should be embedded in any data system to protect its resilience:

- Conducting stress tests focused on potential risks, checking readiness for a range of emergency and non-emergency situations
- Future proofing to anticipate the likely challenges ahead and to ensure data is available on relevant topics
- Finding innovative and creative ways to make full use of existing data considering new data sources, such as big data, and anticipating future data needs
- Learning from expert and vetted staff who provide the best expertise available to support learning, development and technical innovation within data systems.

Theme 6: Good governance

Finding 15: Good governance is seen by participants as essential in helping people to trust and engage with data systems. They see the elements of good governance set out in table 1 as a key foundation on which resilient and effective systems rest.

Table 1

Important elements of good governance
Safeguarding: of the security of the data in the system and the rights and safety of people with regard to the onward sharing of their data.
Independence: to ensure unbiased application and enforcement of the rules across the system that institutions that are rooted in it, including the government, cannot provide.
Effective sanctions: to inspire trust that rules, regulations and sanctions are fairly and equitably applied to everyone involved in the data system whether data collection, storage and management, analysis, interpretation and communication.
Oversight, accreditation, monitoring and inspection of data quality as well as professional standards of those collecting and using data. This would include an audit process to ensure consistent analysis of the data resultant from data systems. It would ensure that someone is responsible for verifying data security and privacy systems are in place and working as they should be.

Participants recommend that:

- Governance of data systems is more widely visible to people across society
- There are transparent communications on the actions regulators have taken demonstrating independence from those who manage data systems
- Current regulators are given boosted powers to penalise misuses of data systems
- Support is provided through the governance structures for data systems to employ best practice which flows from non-emergency to emergency situations
- A culture of learning from the before, during and after emergency situations is embedded in data systems
- Data systems are designed with built in oversight, monitoring and inspection of quality
- Those working data systems have appropriate skills and experience to protect the efficacy and trustworthiness of the system
- Reassurances are made to society about data privacy and security across all data systems

Theme 7: Communication and awareness raising

Finding 16: Participants are more accepting of data systems use in emergency than non-emergency situations. They agree that the visibility of data systems, and for the organisations that govern and regulate them, is lacking. They fear this could lead to misunderstandings and for data being collected and not used – which they find unacceptable.

Finding 17: Simple and honest communications about data systems and their management is essential with effective data systems speaking to society to allay fears, build trust and supporting people to see the purpose and value of participating in the system.

Participants recommend that:

- Clear communications are needed on what data is collected in non-emergency situations to inform what data can and should be used in emergency situations
- Communication campaigns are needed to highlight the public benefit which comes from effective data systems
- Efforts should be made to re-frame perceptions of data use so that public benefit is front of mind when people think about data systems
- Public benefit should be used as a lever to ensure societal needs are met through data use in emergency and non-emergency situations.

Participant conundrums, solutions and areas for further research

Several data conundrums emerged during the course of the dialogue which represent apparent stress points in how people think the data system should work to be fully effective and how they feel as individuals about sharing their own data. The main three data conundrums and related solutions are set out in Table 2:

Table 2

Data system conundrums	Potential solutions described by participants
Precise details are required for data quality, comprehensiveness and an accurate picture of society. However, asking for what people might consider to be too much personal data is a problem for participants. They feel it can make people less likely to engage in data systems for non-emergency situations leading them being missed from data which could support them in emergency situations. The lack of engagement being due to people's perception that they are more vulnerable to harm, exposed to risk, or simply inconvenienced if they do engage.	• Undertake further work to raise awareness in society that data is collected and used for public benefit – including as a key element of responding to emergency situations • Ensure public communication on data systems include clear and simple communication on how and when data is de-personalised[6]; and how personal and sensitive data is protected. • Transparently demonstrate what the benefits of data systems are; creating a shift in public awareness towards an understanding that data systems can bring public benefit.
Many participants believe that data should only be collected for a specifically defined purpose, particularly in non-emergency situations. They believe that individuals should only agree to share data based on this purpose. However, participants also recognise that a resilient data system requires data to meet future needs that aren't yet known. This creates a dilemma – how do you state a clear purpose for data collection when you are not yet clear what these future needs might be?	• Clarify, in simple terms, across a range of emergency and non-emergency situations what the purpose of any given data system is including: • how data collected by private and public sector data systems is used -and why • where data might be shared, and who with – and why • where there may be overlaps between the private and public sector in who 'manages' and 'owns' the data. • Create simple, visual and Plain English/ Easy Read terms and conditions documents for websites and apps which collect data. Which might include colour coding to indicate when specific types of data are being collected e.g. location or personal data. • Create a series of good news stories around data use, e.g. in handling an emergency situation, so that people can see the wider public benefits that can accrue from data systems.

---

[6] Participants found Understanding Patient Data's Identifiability Demystified handout helpful in this context

Table 2 continued

Data system conundrums	Potential solutions described by participants
Participants call for data systems to be more joined up, particularly in public health emergency and non-emergency situations. They believe this will make them more efficient, resilient and accessible. Despite this belief they are concerned that if data is shared across systems, and with all those who need it, this may increase the chances of harms to individuals and make it difficult for people to feel in control of who has access to their data and for what purpose.	• If data systems are to be more inter-operable and linked, then protections must be put in place and communicated widely • If data systems are demonstrated to be resilient e.g. to be able to recover from challenges and adapt to changing circumstances, participants believe people will be reassured that harms and risks have been minimised • Participants feel that joined up systems should prioritise vulnerable people, particularly in emergency situations – using the fact that they are joined up to understand who is most at risk in an emergency • Design data systems with inclusion and diversity in mind including: — Involving a diversity of people in the design of data systems — Standardising the design of systems, particularly those in the public sector so that it is easier to move from one to another. — Having dedicated and specialist teams responsible for system accessibility.

As a result of identifying these conundrums and potential solutions participants a number of areas for further research and future lines of enquiry are indicated, mostly focused on involving people across society in data system decisions. These include:

- Researching ways in which trust in data systems at a local level can be fostered
- Governance structures developing systems, including public involvement panels, which encourage data systems to operate as learning systems
- Studying the facets of trust explored in this dialogue further with a citizens' jury or similar deliberative panel which brings people together over time to test specific data systems against these elements.

Participants recommend that:

- Data systems are shaped, challenged and developed with the involvement of a diversity of people from across society
- Public involvement should inform how data is collected, including the inclusion of data from those who might be missed from the system
- Public involvement should be a key part of data system governance structures

We end this report with a call to action voiced by one participant highlighting the views of many in the dialogue:

> *People will feel like their opinions are heard and it's trusted. Because it's like, Okay, no, we were a part of this decision. We helped make this decision. It doesn't feel like it's being imposed upon us.*
>
> **Participant, UK**

Hopkins Van Mil report authors:

Henrietta Hopkins, Director
Suzannah Kinsella, Senior Associate
Dawn Snape, Senior Associate
Hally Ingram, Senior Associate

Report design by vincentdesign.co.uk

Hopkins Van Mil

Contact
Coppergate House
10 Whites Row
London E1 7NF

Email
info@hopkinsvanmil.co.uk
Visit
www.hopkinsvanmil.co.uk