Royal Statistical Society & Centre for Public Data Improving Data Access Workshop

March 2026

Contents

Background

The Royal Statistical Society (RSS) and Centre for Public Data (CfPD) hosted a workshop on 5 February, 2026 that brought together data users, producers, and systemic thinkers to discuss how to improve access to poverty-related data. This workshop was part of our ongoing research into poverty data gaps in the UK, funded by the Joseph Rowntree Foundation Insight Infrastructure team.

In January, we published an interim report that highlighted specific and systemic gaps affecting poverty research, drawing on desk research, interviews, and roundtables that formed the first phase of our research. A key finding was that barriers to accessing data were consistently a problem for poverty research.

Researchers from the third sector, private sector, and wider civil society face particularly significant challenges with data access compared to academic data users. We therefore focused the workshop on improving data access for those users, whose needs tend to receive less attention than academics. Instead of focussing merely on identifying problems, we sought to collectively consider possible solutions and their perceived feasibility and impact.

This write-up synthesises the various contributions made by attendees, who represented diverse data users from charities as well as producers of poverty-relevant data. We present challenges that were discussed alongside relevant solutions that were brought up that have potential to help overcome them.

Challenges

Friction Points in Data Access

Any process for accessing data will involve some degree of friction when users try to interface with it. This means that data users will have to expend some amount of their own time or resources trying to access the data in the way that suits their needs. Yet, there are varying levels of friction that can affect users' interactions with the system. One goal for the UK's stats producers should be to minimise those frictions as much as possible within the boundaries of accepted security frameworks.

The topic of access frictions must be situated in the context of the technical and resource capacities of the data users in question. These are by no means uniform across civil society. Indeed, our workshop included representative users from diverse backgrounds in this regard. However, there are some general challenges that came through in the conversation.

Many organisations simply lack the skills or capacity to engage with data in meaningful ways. Many of the groups that face the most extreme version of this problem are those that are at the frontline of poverty alleviation efforts—often small teams of people whose roles may not even primarily be related to data. As one attendee who works closely with those users put it, "if you need more than Excel to use a dataset, then [for these users] it may as well not exist."

For larger or better resourced organisations that can afford dedicated analytical teams, the issue is less that they cannot use data in technically sophisticated ways. Rather, they face constraints on their time and capacity that make frictions in accessing data sometimes problematic for their work. Regardless of an organisation's capabilities, examples of points of friction can include, among others:

- Complicated processes for accessing data, or a lack of adequate support for people trying to get and use data.

- A need to navigate multiple datasets, potentially from multiple data owners and multiple access routes.

- Lack of standardisation between regions or datasets.

One salient example of friction in the system is the discoverability of public data. There appears to be a considerable knowledge gap about what exists and how people can access it. For instance, although the systems that users rely on for public data like Stat-Xplore can be useful, some users find it challenging to know which datasets they should look to for the information they want, and better documentation or guidance in this respect would be helpful.

Participants emphasised that the challenge is not simply locating datasets, but also understanding what they contain, how they can be used, and whether they are appropriate for answering specific policy or service-delivery questions. For many users, documentation is written with academic or professional researchers in mind, which can make it difficult to interpret variable definitions, coverage, or limitations. This creates a situation where datasets may technically be available, but remain effectively inaccessible to those without specialist expertise or the time to interrogate them in detail.

Yet, even for experienced researchers, documentation sometimes does not exist or is inconsistently updated. It may also be a challenge to get in contact with relevant people within departments who have expert knowledge of the dataset and the capacity to support on queries.

Taken together, these contributions point to a broader navigation and comprehension barrier within the current data landscape. In practice, the consequence of this gap means that organisations may spend significant time and resources attempting to locate the right data, or may duplicate effort by collecting new data when existing sources could have sufficed. Poor discoverability of data is therefore a friction point that directly affects efficiency.

Of course, producers also face resource constraints, so investments to reduce friction have opportunity costs and they need to prioritise wisely. However, many users currently rely on FOI requests and Parliamentary questions (PQs) when they fail to find data through public avenues. Investments that reduce FOIs and PQs may therefore be offset by freeing up resources that are usually dedicated to responding to them.

Permissive Gateways and Access Controls

Producers are understandably concerned to safeguard against unauthorised release of identifiable data. They have certain legal and ethical responsibilities to secure data, and there are some restrictions placed on the kinds of data access routes that producers can open. Release of data to Trusted Research Environments (TREs) relies on the legal permissive gateway provided in the Digital Economy Act 2017, but other access arrangements may not have as clear legal justification.

Moreover, some attendees at our workshop flagged that producers have practical concerns about releasing even non-identifiable data if they might be misused or misinterpreted. This risk is perceived to be greater when released data are not adequately checked for quality or robustness.

Yet while the reasons for prioritising security are fair, there are trade-offs that need to be considered between stringent access controls and the usability and impact of data. To illustrate this point, consider the case of frontline anti-poverty charities. They typically need to be able to identify specific populations for their analyses, often defined by specific combinations of demographic characteristics and specific geographies. Doing so is important for robustly assessing the impact of their interventions or intelligently targeting their services in new areas.

As we've heard throughout our research, these sorts of projects are often inadequately served by publicly available data. For these users, the challenge is often one or two missing characteristics necessary for identifying or analysing their target groups. Typically, it is not that the combination of characteristics is strictly absent from existing datasets, but that those combinations could—in some cases—produce potentially identifiable data and are therefore suppressed from public versions.

Of course, needing to access anything other than public datasets is a significant barrier for many users who may not have time or resources to do so. For researchers who do have the capacity to engage with the established data infrastructure, formal access processes can be slow and opaque. Sometimes this means that projects simply cannot happen given "unrealistic timescales for approval." The burden on researchers time is also a recurring cost of working with data, since applications typically have to be submitted repeatedly for each project, even if they are simply running a previously approved analysis on a new release of the same dataset.

Transparency, Incentives, and Trust

One challenge that affects users across the spectrum is a general lack of transparency around how decisions are made. Attendees flagged that they find communication lacking about why some data are held in secure environments instead of the UK Data Service, or why their FOI requests are rejected, for example. They also reported that they did not think there were clear "escalation" routes when engagement with producers falls short or they are not given clear justifications for rejections.

As discussed above, producers are (understandably) risk-averse in their interpretations of current legal frameworks when making decisions about releasing data. Indeed, their credibility as institutions is partially tied to their handling of data securely. At the same time, lack of transparency about access decisions is an important issue that affects users' trust in the system. Here are some of the written contributions we received during the session about this topic:

- "Secure access requirements [are] used to obscure data."

- "Data completeness – hiding [valuable data] behind 'protected sensitive data' GDPR excuses."

- "[There is a] lack of motivation to improve data access by data producers."

- "Lack of data – redaction hides other data."

- "IT dept. default is [to] keep things 'in house' + protected."

This kind of language reflects a certain degree of cynicism towards the current system's justifications for restricting data access. There is perceived overuse of legal compliance to restrict sharing for what are sometimes considered political

reasons. Interestingly, one attendee noted that, in their experience, challenges with getting and using data can lead people to give up on taking data-driven or quantitative approaches to their work entirely.

Maintaining trust in the system is therefore not just about ensuring the security of data, but also about transparently communicating why decisions are made and what recourse people have if they think access is being unnecessarily restricted.

Potential Solutions and Areas for Further Exploration

Participants proposed a range of potential solutions to address the challenges identified during the workshop. While these ideas varied in their level of specificity and feasibility, several coherent themes emerged that will help shape our thinking as we move into the final stages of this project.

Taken together, these proposals suggest that improving data access for non-academic researchers will require coordinated action across multiple fronts: new intermediary mechanisms to bridge gaps between producers, funders, and users, stronger networking and knowledge-sharing structures including through explicit funding structures to support community involvement, improved discovery tools and training, and more proportionate, transparent access processes. While not all of these ideas can be implemented immediately, they offer a useful starting point for future policy and system-level development.

Developing an Intermediary "Middle Layer" Between Producers and Users

One of the most prominent ideas was the creation of an intermediate layer between statistics producers and the diverse communities of data users. Participants suggested that many of the current frictions arise from the absence of a clear, trusted interface through which user needs can be communicated and access processes navigated. Although the discussion could not go into detail regarding how this would work, a few potential models were suggested by attendee contributions.

One model could be an intermediate layer that would take the form of a convener, "data broker" or similar body that would support user groups in identifying relevant datasets, understanding access routes, and formulating requests in ways that align with producers' requirements. At the same time, such a body could help aggregate and elevate user needs, presenting them to producers in a more structured and manageable way that makes priorities clearer.

An alternative model might involve the development of trusted intermediaries capable of undertaking analysis on sensitive or linked datasets on behalf of third-sector organisations that lack the capacity to apply directly for access themselves. In this model, analysis could be conducted within secure environments by trusted research organisations—perhaps reputable thinktanks—with results shared back to the requesting organisations.

Overall, these ideas reflect a recognition that the current system often assumes a level of technical capacity and institutional familiarity that many users do not possess. An intermediary layer could help bridge this gap while maintaining appropriate safeguards around sensitive data.

Strengthening Networking, Dialogue, and Resource Sharing

A second cluster of proposals focused on improving communication and knowledge-sharing across the data ecosystem. Participants repeatedly emphasised the value of creating more opportunities for two-way dialogue between producers and users, a point that was summarised in the workshop simply as a desire for "more chat."

This could include regular forums that bring together statistics producers, third-sector organisations, and other users to discuss emerging needs, challenges with specific datasets or tools, and planned changes to data releases. Such spaces were seen as a way to build mutual understanding, reduce misperceptions about constraints on both sides, and ensure that data production is better aligned with real-world needs of policy and service delivery.

At the same time, some users emphasised the need to allow for more abstract, conceptual conversations with producers about how to understand poverty. Many of those topics would benefit from the contributions of people affected by poverty, although there is a need to translate their issues and experiences into specific, data-centric points that will be useful to and actionable by data producers.

This may be another space where an "intermediary" layer of communication between producers and the wider user community may be helpful, as it could help facilitate these conversations and translate issues into specific requests about specific data wishes. It could also help produce venues where there is room to debate more philosophical or systemic decisions (i.e. the ONS's prioritisation plans) in a way that is sensitive to producers' need to be politically neutral— something that some attendees did not feel is well-served by existing channels of communication.

In addition to formal forums, participants suggested the development of more user-friendly interfaces and platforms for sharing best practice, teaching resources, and practical guidance on using key datasets. Networking mechanisms and resource-sharing platforms could help disseminate knowledge about what works, reduce duplication of effort, and support organisations that are newer to working with data.

Improving Tools and Investing in Capacity

A third set of proposals centred on improving the tools available to users and investing in their capacity to engage with data effectively. One idea was the development of a centralised catalogue or discovery tool that would clearly explain what datasets exist, what information they contain, and how they can be accessed. Participants envisioned a single, searchable "data library" interface covering major UK administrative and survey datasets, designed explicitly with non-academic users in mind.

Such a tool could provide compact, comprehensible summaries of key datasets, including information on coverage, available variables, possible disaggregations, and access routes. Some attendees also suggested that more intelligent search or navigation tools—potentially including AI-assisted approaches—could help users identify relevant datasets and variables more quickly, particularly where they lack detailed prior knowledge of the statistical system.

Alongside improved tools, participants emphasised the need for additional investment in training and support for non-academic users. This could include both training in basic data skills and more advanced guidance on navigating secure access processes, interpreting documentation, and linking datasets. Within academic contexts, this could also be woven into

more binding involvement of civil society organisations in meaningful co-production and research development, potentially supported by explicit commitments from funders. Strengthening user capacity was broadly seen as a valuable complement to any reforms aimed at improving access itself.

Reforming Access Processes and Governance Norms

Finally, participants proposed a number of reforms to the way access processes and governance norms currently operate. These included the introduction of fast-track application routes for researchers or organisations with secure setups or a strong record of compliant and responsible data handling. Such pathways could reduce delays for low-risk users while allowing producers to focus scrutiny where it is most needed.

Closely related was the idea of more proportionate, tiered access pathways that reflect different levels of risk and organisational capability. Rather than a one-size-fits-all model, access arrangements could be calibrated to the sensitivity of the data and the demonstrated capacity of the requesting organisation.

Participants also discussed the need for broader cultural change in how data access is approached, particularly in relation to FOI requests. A stronger "default to share" culture—balanced with appropriate safeguards—was seen as a way to reduce reliance on formal FOI processes, which are often time-consuming for both users and producers. As more data becomes proactively available through platforms such as Stat-Xplore and similar tools, the need for ad hoc requests may diminish, leading to a more transparent and efficient overall system.

Naturally, cultural changes of this sort may have to be accompanied by expanded provisions for legal permissive gateways in legislation to be fully effective. However, as noted above, there are trust-related challenges that changes in culture or disposition can positively affect. It should be a priority to move towards a system where users from across the spectrum of capacity and resource can trust that their interests are being looked after.

Next Steps and How to Get Involved

We are in the process of synthesising findings from across our various workstreams and will produce a set of clear recommendations for fixing certain data gaps, which will be targeted at multiple stakeholder groups. The solutions that were proposed at this workshop will help inform those recommendations and our advocacy work going forward.

The work from this project is set to conclude in March 2026. One of our last workstreams will develop case studies of specific data gaps and their consequences that will help illustrate the importance of finding solutions to some of the issues discussed here. If you have specific data gaps that have affected you that you think we should be aware of, please email Dakota Langhals, RSS Policy Researcher, at d.langhals@rss.org.uk, or policy@rss.org.uk.

We also recently launched our Poverty Data Gaps Explorer tool. The intention is for it to grow into a crowd-sourced catalogue of data gaps that can help organise and coordinate advocacy for the fixing of specific, data-related issues in poverty research. Currently, you can use it to explore gaps that the project team extracted from hundreds of civil society reports about poverty that were published since 2020. We also invite you to submit your own gaps to help grow the tool into a more dynamic, user-driven resource.