Future Data Services
Data Access Theme Paper 4
Home and remote access

Author: Felix Ritchie, Senior Strategic Fellow to FDS.

August 2024

Future Data Services is a two-year review by the ESRC into the operations and management of UK data services. It has five themes: data discovery and curation; data access, user support and training; technology; people, organisations and culture; and ethics, public engagement, and impact. This piece is for the **data access** theme.

Papers produced under this theme are reflections on the evidence gained during the review, augmented by practical experience and familiarity with the literature. They are intended to challenge conventional perspectives and propose new ideas or ways of working. They provide the arguments behind the recommendations of this theme.

The views expressed in this paper do not necessarily represent those of ESRC or the Future Data Services Project.

Reference title: FDS (access theme) Paper 4 *Home and remote access*

Contents

Summary of recommendations

discussed in this paper

Recommendations for data services

A4.1 Confirm home working as the default position for access, with requests for restrictions to be considered in accordance with guidelines

Recommendations for ESRC

A4.2 Demonstrate the evidence base on researcher trust and efficiency gains to all parties

A4.3 Develop good practice/good principles for remote working, including connection types, acceptable locations, what needs to be in organisational remote working policies

A4.4 Explore the drivers behind limited international access and develop potential solutions

1. Issue

Prior to the pandemic access routes to confidential microdata were primarily either safe room (for government facilities) or remote working (for academic organisations). The pandemic shifted perceptions of what was possible, secure and desirable. The movement towards remote working (access from the researcher's institution) and home working (access from private locations) has greatly increased research productivity, with no evidence for an increase in risk.

There is variety in practice, but also a considerable amount of agreement, so that it is feasible to draw up guidelines on good practice, as done below. However, development is held back by institutional factors: lack of commitment to a long term vision, guidelines which lag behind practice and/or do not reflect reality, and a default-closed perspective on the part of data providers which creates uncertainty.

This short paper reviews the options for home and remote working in various data facilities in the UK as they stand, areas of commonly-understood good practice, and proposals for future good practice. The evidence was gathered from interviews with secure facilities in 2022[1] and with users and data professionals throughout the FDS project.

We propose a move towards a default-open principle, where home working is the default access for accredited TREs, along with recommendations on what constitutes 'good practice'. We also identify some initial first steps.

Definitions used in this paper:

- "safe room": access from the service's premises or an accredited safe room such as the SafePods, UKDA safe room, NISRA or Scottish Government access rooms
- "remote working": access from one's institution, rather than a safe room
- "home working": access from a non-institutional location
- "IAO": Information Asset Owner, the person responsible for data release decisions

[1] Details can be found in the appendix to Ritchie F. (2023) Home and Remote Access. Report for the Future Data Services project.

- “Organisation”: the researcher’s employer eg university or government department
- “RDC”: research data centre, a TRE allowing users to have ‘virtual desktops’ and work as if the data were on their own machines
- “RJS”: remote job server, allowing users to send code and get statistical results back; currently this only applies to OpenSafely in the UK

2. Current practices and perspectives

2.1 Pre-pandemic and pandemic positions

Most of the general-purpose social science TREs in the UK (eDRIS, ONS, SAIL, UKDS), plus the NHS TRE and OpenSafely), offer a remote or home working solution. Only HMRC and NISRA still require visits to a physical secure room. SAIL, UKDS, OpenSafely and the NHS TRE offered remote access from startup, as this was part of their function (although OpenSafely and the NHS TRE only began operations in the pandemic). ONS had begun offering remote access before the pandemic. During the pandemic it switched to remote access and closed its safe room; eDRIS did the same.

For all organisations offering remote access, home access became an additional option during the pandemic. To date, all organisations have retained it. For organisations who have changed their access policy during the pandemic, there do not appear to be explicit statements on the rebalancing of risks and controls created by the new ways of working. ONS’ position, for example, has been made informally on a number of occasions but is not yet stated in a risk review.

2.2 Views on remote/home working

Positive benefits from home working included:

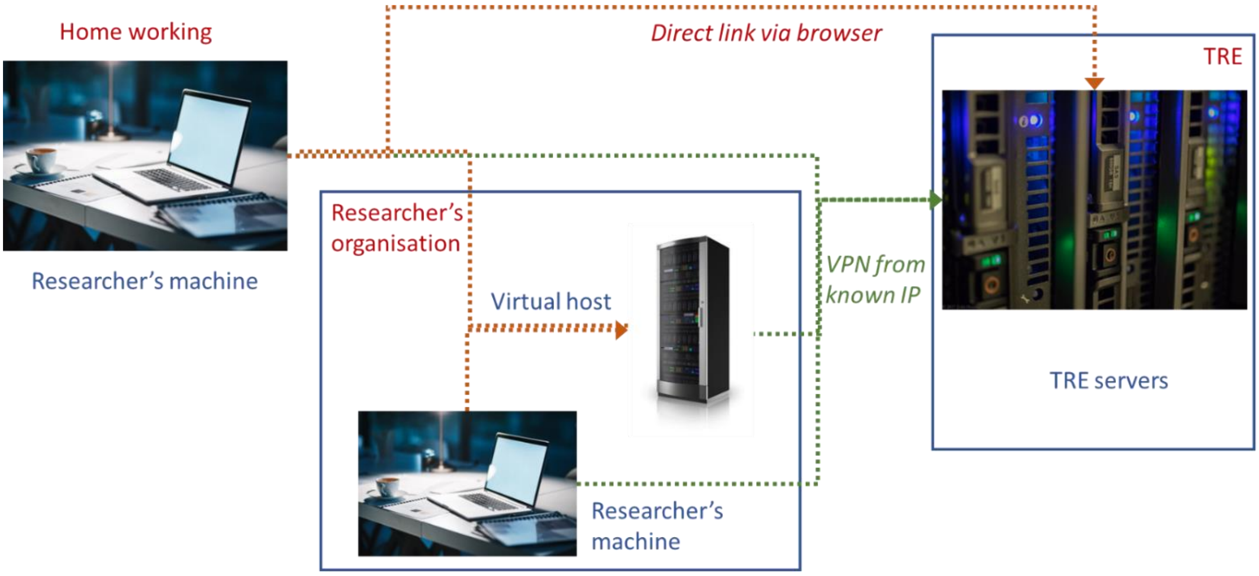
- Better experience for researchers: less travel, more time flexibility, screen sharing
- Fewer outputs requested
- More inclusive – supporting of disabled or time-constrained researchers

No negative views on home working were expressed. It was acknowledged that technically home working allows less control over where the research operates from. However, there was common agreement that researchers generally act appropriately and follow instructions, and that no risks to confidentiality have been reported. ONS explicitly balance the theoretically increased ‘safe setting’ risk (working in an inappropriate place) with an increased ‘safe people’ control (additional commitments from researchers’ employers).

There is a recognition that data providers/IAOs do not identify the benefits and are keen to restrict access to remote access or safe rooms. For ONS, eDRIS and OpenSafely the current arrangements are, strictly speaking, temporary, although only the Scottish Government had stated an intention to return to safe rooms.

2.3 Technicalities of linkage

Remote and home working is usually delivered via an organisational VPN:



A virtual private network (VPN) is used to connect the organisations to the TRE servers. This creates a 'tunnel' preventing the unauthorised views of the researcher's screen. The IP addresses (identifiers) of the organisation computers are generally whitelisted ie only known IP are allowed to connect. Remote working is achieved by one of two methods:

- A 'virtual host' in the organisation acts as the link to the TRE via a corporate VPN, with a whitelisted IP address
- The researcher's computer has a whitlelisted IP address and connects directly to the TRE via a corporate VPN

There is more variation in home working. Some organisations require access through the organisation's virtual host; connection to the virtual host may be via VPN or an unencrypted connection. Alternatively, the researcher may be able to directly connect to the TRE, perhaps but not necessarily via a VPN. IP addresses are not whitelisted in this case.

TREs provide a view into a research environment, but not the data itself (OpenSafely does not provide a view either, but allows code to be sent and results returned). The VPN link means that the view from outside the TRE establishment is encrypted. A second VPN link means that there is double encryption. A direct link to the server from the researcher's machine means that the IP of the researcher can be checked (eg to see whether the researcher is physically in the UK, assuming no IP impersonation), but makes whitelisting harder. Access via a virtual host means only that the IP address of the host is known, not the end user. The technical connection therefore embodies a lot of choices, which may not have been explicitly made by those organisations.

2.4 SafePods

'SafePods' are an approved form of safe room, with 21 deployed at universities around the country. This enables researchers who do not have a home working or remote connection, or who are blocked by data owners from home working, to access the social science data TREs, all of whom allow access through SafePods. This therefore widens significantly access options for some researchers, although access is limited to working hours (10-4pm Mon-Fri) and to the availability of the SafePods, which can be booked up weeks in advance.

SafePods are metal boxes, roughly 1/3 of the size of a shipping container and of similar appearance. They have biometric door locks, CCTV, an internet connection, large desktop screens, and a telephone to contact the SafePod help desk. Screen sharing is not possible with SafePods.

The physical size of the SafePods is a problem for the host organisation (IT problems are minimal), and so the next development is 'SafePoints'. These effectively take over existing office space, setting up secure terminals and installing a door lock which communicates with the SafePod admin team to regulate entry. As well as providing a better environment for the researchers, this is more appealing to the university as it re-uses existing space, and there is flexibility to allow multi-use rooms as long as the room is not being used by SafePoint researchers at the same time as other staff.

In theory home working should lead to substantial reduction in demand for the SafePods, but this does not seem to be occurring at the moment. Two factors contribute to this: IAOs insistence on access through a safe room, and organisations not setting up their own remote connection.

2.5 Researcher perspectives

Researchers were not formally interviewed as part of this review. However, they have made their views clear to the research team in other circumstances. The research community recognises the value of remote or home access. There are substantial productivity gains from

- Not having to travel to physical sites
- Ability to work on-and-off for short periods pf time
- Working outside of office hours
- Screen sharing with other members of the project team

Remote working has been very important but home working has been the big (and positive) step change in access.

Researchers have raised a number of concerns:

- Academic organisations have been slow to commit to the agreements for remote working, particular ONS' Assured Organisational Connectivity (AOC)
- The need for university laptops to access remote/home working solutions is a restriction, particularly as universities may have hardware replacement cycles that do not reflect research needs
- SafePod geographical distribution is patchy, and the high demand makes it hard to plan research time
- Applying for home working on project-by project basis for IAO approval is time-consuming and adds uncertainty
- The closure of the ONS safe room in London has dramatically reduced access options in the capital

3. Common features and gaps

3.1 Agreement and disagreement

For remote working there are only two unanimous requirements for remote working

- There needs to be an organisational level agreement with the TRE
- Access must be via a machine supplied and maintained by the researcher's organisation

However, for most organisations there is also agreement that for remote working

- Access is through an organisational VPN
- IP addresses are whitelisted

For home working there is more variation but all agree that

- The organisation must have an acceptable home working policy, which includes guidance on not working on confidential material in public areas

And most agree that

- Access should be via a direct or indirect VPN connection, at least for the final link to the TRE

There are also areas where there is agreement but the position itself is not formally started or is unclear

- All non-health TREs agreed that the IAO has a veto over where data can be accessed from; but it is not clear what the default expectation is and whether IAOs have to justify more restrictive access or TREs have to argue for more flexible access
- All RDCs allowed screen sharing between accredited researchers but none had a formal policy on it, and there is some confusion amongst users (and service providers)

3.2 Screen sharing

On screen sharing, in 2022 only the nascent Integrated Data Service allowed for it in its draft user agreements. The theoretical concern is that a researcher might share screens with those not authorised to access the data; this may be why TREs are reluctant to put this in their protocols, arguing that 'data owners' (and others unfamiliar with TREs) won't accept it.

There is no evidence to support this contention that researchers cannot be trusted on this. On the contrary, researchers appreciate the value of screen sharing, are conscious of the risk to their research if it were removed, and so actively protect it.

There is an expectation that researchers would do it anyway even if it were explicitly banned, as it massively increases productivity; hence, while not explicitly approving it, TREs were not foolish enough to explicitly ban something of substantial research value but which they couldn't monitor or control. As a result, in 2024, the 'don't ask, don't tell' approach still seems to be dominant.

Screen sharing has significant advantages for the TRE by ensuring work stays within the TRE and significantly reducing output checks. For example, conference presentations can be developed jointly within the TRE, or complex research findings discussed before deciding on the final versions to be released. For those organisations requiring a two-stage release procedure (intermediate and final), screen sharing allows the intermediate stage to be scrapped without significantly affecting research.

3.3 Where is 'home'?

It is not clear what counts as 'home' working. Some organisations require the researcher to specify a particular 'home' address, and implicitly or explicitly only allow access from that location (rules-based model). Others are more concerned to make sure that the access point is secure: that is, not in a public place, and not using a public internet to connect (principles-based model).

There is no evidence on how well individuals stick to these restrictions, but the rules-based model is likely to be far more problematic. It means that, for example, an academic visiting a different university would not be able to access the facility even if they were securely logged on using EduRoam; nor would they be able to use a secure mobile hotspot from a hotel room. These restrictions are likely to seem excessively cautious and so may well be ignored. The principles-based approach is more likely to achieve buy-in, partly because it is more flexible, and partly because it is likely to be consistent with the researcher's employer's remote working policy.

3.4 International access

Few organisations formally allow international remote access or homeworking except in specific cases, but most agree that a GDPR-type data regime in the organisation's location is likely to be a precondition.

4. Taking a strategic perspective

4.1 First principles

The UK's various agreements have grown up piecemeal, some as part of long-term visions, others as a response to circumstance, especially the pandemic. Agreeing a core set of principles to be adopted as a default position by TREs would simplify the landscape, clarify procedures and provide strategic direction for organisations not yet working in line with those principles.

Like other Future Data Services analysis, this follows the EDRU (evidence-based, default-open, risk-managed, user-centred) model of data governance. In this case, the key pieces of evidence are the demonstrable ability of accredited TREs to provide home and remote working safely, and the substantial productivity gains that have resulted.

The basic principle proposed is:

> *Home working within the UK should be considered as the default position for an accredited TREs; data providers requiring additional restrictions (remote working, safe rooms/SafePods) should be required to provide evidence for the exception*

4.2 Sharing good practice

Good practice can be defined as:

- Server access via organisational VPN
- Connection via organisational machine
- The user organisation should be able to demonstrate sensible home working policy about access sensitive data in public areas
- Clear statements on
    - Acceptability of 'non-home' location (eg working in hotel room, working at other organisations, working temporarily from location abroad)
    - Guidelines on working practice to be followed in those locations (eg not being overlooked; no public wi-fi)
    - International access
    - Screen sharing

4.3 Practical first steps

Getting agreement that IAOs having to make the case for a more stricter environment than the default is a reversal of current practice, but this is what current policy is supposed to be. Experience in the past suggests that IAOs quickly become habituated to social norms; this why a default-open principle is so important. Opening these conversations with data owners is an important first step in changing perspectives.

A surprising element of the analysis is why AOC is not almost universal, when it offers significant benefits to researchers for relatively little organisational commitment. There is a perception that this is difficult to achieve, and in the early days of AOC the form was unnecessarily complex for academic user. However, our experience is that this is largely driven by institutional issues at the university. Improved guidance (ideally with the support of both ONS and ESRC) could address this.

Safe Researcher Training and guidelines need to be reviewed and updated to reflect home working. TREs may also wish to review their output checking model.

Recommendations for data services

A4.1 Confirm home working as the default position for access, with requests for restrictions to be considered in accordance with guidelines

Recommendations for ESRC

A4.2 Demonstrate the evidence base on researcher trust and efficiency gains to all parties

A4.3 Develop good practice/good principles for remote working, including connection types, acceptable locations, what needs to be in organisational remote working policies

A4.4 Explore the drivers behind limited international access and develop potential solutions.