UK Data Service

Research Data Handling and Security: Guide for Users

Public

17 November 2020
Version: 08.00

T +44 (0)1206 872832
E collections@ukdataservice.ac.uk
www.ukdataservice.ac.uk

Contents

Scope

This guide is for users of research data accessed via the UK Data Service (the Service) through its online services provided by the UK Data Service. In particular, all users who obtain Special Licence data or Secure Access data, where disclosure risk is increased, are required to read and understand this document under the terms and conditions of access.

In this document we use the terms *Safeguarded* to describe data which is made available under the End User Licence, and *Controlled* to describe data which is defined as personal under the Data Protection Act (2018) or the General Data Protection Regulation and is made available through secure mechanisms.

1. Licence framework

The Service does not own the data held in its collection but it is licensed by the data owners to curate and share the data on their behalf. Users accessing the data have responsibilities to preserve data confidentiality and to observe the ethical and legal obligations pertaining to the data. In particular, users must maintain the commitments made to survey respondents to preserve the confidentiality of the data provided.

The conditions under which data may be accessed are specified in a deposit licence and are also set out clearly in the catalogue record's Access Conditions.

Importantly, these conditions include providing the data only to users who have registered with the Service's

online services and agreed to an End User Licence (EUL).

1.1. End User Licence data

Use of Safeguarded data is governed by a legally-binding EUL which forms part of the registration process. Each individual who requires access to data has to register with the UK Data Service and will need a UK Access Management Federation (UKAMF) login. Users who are part of the UK Higher/Further Education (UK HE/FE) sector will automatically be issued with these details by their organisation. Users who have no other way of obtaining a UKAMF login can apply to the Service.

Under the terms of the EUL, users agree:

- not to use the data for any commercial purpose (except with prior permission/under an appropriate commercial licence agreement);
- to preserve the confidentiality of, and not attempt to identify, individuals, households or organisations in the data;
- to use the citation and acknowledgement information provided by the Service, in publications;
- to supply to the Service, the bibliographic details of any published work based on the data collections;
- to ensure that the means of access to the data (such as passwords) are kept secure and not disclosed to anyone else;
- to abide by any further special conditions.

1.2. Additional conditions

For some data collections, additional conditions to those of the EUL may be specified by the data owners. This may or may not relate to disclosure risk. Typically Special Licence data collections have a higher of risk of disclosure, but are not considered as personal data.

Users are prompted within their User accounts to sign and agree to any additional conditions, and follow any steps required for accessing data, such as gaining explicit permission from the data owners. Access to Special Licence data may be restricted to certain users, for example, those working within EEA Adequacy countries, who follow the GDPR regarding data handling and storage.

Data explicitly defined as **personal** cannot be accessed via a Special Licence. These **Controlled** (Secure Access) data are only available through the UK Data Service secure environment, the Secure Lab. Any registered user requiring access to secure data will have to (i) be accredited as an ESRC and/or UK Statistics Authority Accredited Researcher, (ii) complete training and pass test and (iii) agree to the Secure Access User Agreement.

2. Accessing data

Data can only be accessed under certain conditions:

- under the EUL, data can only be accessed by registered users;
- data supplied under additional conditions can only be accessed by those who have accepted these conditions;
- Special Licence and Secure Access data can only be accessed by approved individuals for a specified usage and for a specified time;
- Secure Access data can only be analysed remotely within the Service's Secure Lab and outputs are only released to users subject to statistical disclosure control by Service staff.

2.1. Research projects and teams

Users are required to register research projects via their online UK Data Service account. A research project can comprise a single user or project team. The period of access is typically set by the data owner for Special Licence and Secure Access data, but tends to be between 2 and 5 years. Users should contact the UK Data Service Helpdesk if they wish to extend a project.

Where a user joins a research team that is using Special Licence or Secure Access data:

- the new user must add the data to their project and follow any workflow step including completing necessary applications forms and additional user agreements;
- approval must be sought from data owners and gained before the user can access the specific data;
- the user must complete SRT training and pass the SRT test to access Secure Access data;
- the Service will provide advice on the process to be followed, which is largely self-prompted within the User Account system.

2.2. Teaching purposes

When using data for teaching purposes, the tutor must sign and return the Teaching Agreement ad ensure that all students are registered with the UK Data Service.

Special Licence and Secure Access data cannot be used for teaching.

2.3. Re-use of data

To re-use anything but open data already supplied, but for a different purpose, it is necessary to re-apply for access. All data usages are assigned to a dedicated project in a User's account. Any new project will need to be registered and data added. Where a data owner's permission is required, this will need to be obtained again for a new project. To use data for a different project is a breach of the User Agreement and subject to penalties.

2.4. Security

Passwords and passphrases that provide access to UK Data Service data or computing environments must never be disclosed to anyone else. Data must never be stored on a computer that might enable unauthorised access.

3. Data storage and security

3.1. End User Licence data

All data provided by the Service must be stored under conditions that meet the undertakings given in the EUL (see section 7 for organisational responsibilities) and those listed below. Additionally users should be aware of any **specific information security guidelines** provided by their organisation.

Authentication

- access to PCs on which research data supplied by the Service are held must have suitable personal authentication (i.e. protected by a username and password (see section 3.4 for details));
- if data are placed in a shared location, access must only be available via personal authentication, to those permitted to use the data;

Encryption/passwords

- means of access to the data (such as passwords or passphrases) must be kept secure;
- data on portable media (e.g. a back-up on CD or a USB memory stick) must be encrypted using a secure password/passphrase;

Data deletion

- data must **be deleted** upon project completion as set out in section 6.1. Not deleting data or going on to use it for a different project constitutes a breach of the EUL

3.2. Special Licence data

In addition to the responsibilities under the EUL in 3.1, Special Licence data have some specific requirements around safe data storage and handling. Where a Data Management Plan has been created for a research project, it is useful to refer to any processes agreed with your organisation, and to include these in the project application. Special Licence data:

Setting

- must only be accessed **in an organisational setting**, via an endpoint device (e.g., PC, laptop or thin client) on a closely controlled LAN with restricted access and must not be accessed at a private

residence. Working at home is not permitted unless explicit permission has been applied for, via the Service;
- must be stored in physically secure conditions (e.g. any portable or printed copies must be stored in a locked cabinet with restricted access);
- must be held on an endpoint device that is in a room that is locked when unattended;
- must have access appropriately restricted where stored on a cloud-based service. The service must store data in the UK or an EEA Adequacy-based data centre, and access must only be available via personal authentication, to those approved to use the data;
- must be accessed on a site which has **security standards** that meet the guidelines in this guide;
- must only be accessed at any additional access requirements regarding site settings set by the data owner.

Encryption/passwords
- must be encrypted when not in use using passphrases instead of passwords;
- must be protected by a screen lock with an interval of no more than fifteen minutes and that requires a secure password to unlock it;

Data deletion
- Special Licence data must be deleted upon project completion as set out in section 6.

3.3. Access to data held within the Secure Lab

The UK Data Service provides two methods of access to confidential data via the Secure Lab:

- remotely from the user's organisation;
- from the Service's safe room, physically located at the UK Data Archive in Colchester;
- via the ESRC Safe Pod Network 'SafePods', installed in some institutions around the UK.

3.3.1. Access via remote access

Data accessed remotely via the Secure Lab must:

- only be accessed from a designated office at an organisational or institutional site. If this office is shared with other people, photographs of the office layout need to be submitted to the Service during the account set up process to ensure that the surroundings do not allow unauthorised people to gain access to or view Secure Access data.
- Only be accessed from an endpoint that:
    - is owned and managed by the institution or organisation from which the Secure Lab will be accessed;
    - has a direct connection to the internet via a wired Ethernet connection;
    - has a dedicated public IP address that is unique to the endpoint;
    - has no other network interfaces connected except for the one being used to access the Secure Lab; this includes using VPNs;
    - is not running any services which allow third parties to connect to the workstation e.g. a web server or email server.

When accessing data via the Secure Lab, it is not technically possible for a user to transfer, download, copy and paste any data, or print to a local computer. Users must not copy screenshots to a local computer.

Outputs from analysis of the data within the Secure Lab are only released to the user subject to statistical disclosure control checks undertaken by Service staff. Users are strictly forbidden from copying anything from the screen by any means. Secure Access data and outputs must not be seen on the user's computer screen by unauthorised individuals.

Users who have been approved to work together on the same project may only share unchecked outputs from that project with each other in the relevant shared project area within the Secure Lab. Temporary or duplicate files should be deleted by the user(s) from the Secure Lab. The Secure Lab User Guide provides advice on managing work in the Secure Lab in some detail.

All user activity within Secure Lab is recorded to provide the Service with information about any suspicious activity.

3.3.2. Access via the UK Data Service's safe room

The conditions under which data are accessed via the UK Data Service's safe room are similar to those for accessing the Secure Lab remotely. However, access via the safe room differs from remote access in that:

- access is only available from within the room;
- thin-client terminals are used to access the Servers where the data are held;
- users must abide by the procedures, listed in the *safe room procedures* document (currently known as CD226-SafeRoomProcedures)*;*
- users are required to visit the UK Data Service to carry out their research, and to undertake a special training programme to ensure they are aware of how to safely use the Secure Lab.

3.3.3. Access via SPN Safe Pods

The conditions under which data are accessed via Safe Pods are similar to that of the safe room.

- Sessions must be booked via the Safe Pod Network
- Users must abide by the Safe Pod Network procedures

3.4. Passwords and passphrases

Passphrases differ from passwords in format and in length. Passphrases are usually much longer, up to 100 characters or more and may contain spaces. The greater length and format of pass-phrases makes them more secure. Wherever possible we expect passphrases to be used.

A passphrase must contain a combination of at least eight alphanumeric and symbolic characters. More is better.

Passphrases must:

- not be disclosed to anyone else;
- not be written down;
- be changed at least every three months;
- not be easily guessable.

The Service will provide users with personal logins to access the Secure Lab. Users are required to change their passphrase on first logon, and then to renew it every three months.

3.5. Audit of confidentiality and security procedures

The Service and the research data owners reserve the right to conduct an onsite audit of the licence holder's confidentiality and security procedures and practices, or to require a report of such an audit. For the purpose of conducting an audit, the Service and the owners reserve the right of entry to the premises where the data are stored and/or accessed. (Also see section 7.1).

4. Statistical disclosure

4.1. Data matching

Data matching can increase the risk of the disclosure of personal information and is therefore only permitted under certain circumstances. This must be declared in applications, where appropriate.

4.1.1. End User Licence data

Where EUL data are matched with external data sources this must not be for the purposes of identification.

4.1.2. Special Licence data

Any plans to match or attempt to match individual or household records to any other data source at the level of the individual or household must be declared in the project application, and can only be undertaken with the explicit permission of the relevant data owners.

4.1.3. Secure data

Whilst the Secure Lab provides an area in which secure data could be linked (e.g. with another dataset in the secure collection or with the user's own data) this is strictly subject to the explicit permission of the data owner. Users will only be able to access those datasets approved for a particular research project. It will not be possible to subsequently add new data without a revised application and subsequent approval.

ONS business data may be linked using the anonymised reference numbers (known as IDBR references). A user may be able to produce a larger 'combined' dataset, with many variables providing characteristics that will directly identify an organisation. While this is an acceptable risk within the confines of the Secure Lab, users must be aware that output requests containing information that will identify an organisation, will be rejected (see Section 4.2).

4.2. Outputs

All outputs resulting from analysis of Special Licence and Secure Access data, by members of an approved project, must be subjected to disclosure control checks and treatment. Users must refer to detailed procedures in the following best practice guidance documents:

- GSS Disclosure control for tables produced from surveys (https://gss.civilservice.gov.uk/policy-store/gssgsr-disclosure-control-guidance-for-tables-produced-from-surveys/
- The Safe Data Professional Group Handbook on Statistical Disclosure Control for Outputs (https://ukdataservice.ac.uk/media/622521/thf_datareport_aw_web.pdf)

There ae some differences between SDC management for Special Licence and Secure Access data. The latter are more sensitive and contain variables that can directly identify survey respondents.

4.2.1. Special Licence data

For Special Licence data some simple rules of thumb apply for ensuring that disclosure is avoided:

1. **Tables that contain very small numbers in some cells may be disclosive**

- Tables must not report numbers or percentages in cells based on only one or two cases. Cells based on one or two cases may be combined with other cells or, where this is not appropriate, reported as zero per cent.

2**. Tables and other outputs must not be published in a form where the level of geography would threaten the confidentiality of the data.**

- To guarantee safety, outputs from Special Licence data should not be published if the geography is lower than UK Government Office Region (GOR);
- If there is a requirement to publish outputs from Special Licence data with a lower level of geography, e.g. between GOR and local authority, then the user must consider whether there is a risk of disclosure;

- Where there is any doubt, the user must contact the UK Data Service via the Helpdesk to obtain confirmation of the acceptability of publication of the output if the geography is below GOR. **No outputs may be published with a geography below local authority.**

3. Care must be taken to ensure that individuals, households or organisations cannot be identified in models or other statistical analysis.

- Results based on very small numbers must be avoided;
- Any output that refers to unit records, e.g. a maximum or minimum value, must be avoided;
- Models must not report actual values for residuals.

4. **Graphical outputs must be based on non-disclosive data**.

- Particular care must be taken not to report extreme outliers.

4.2.2. Secure Access data

The Statistical Disclosure Control (SDC) requirements for Secure Access data differ from those mentioned above for Special Licence data.

Access to Secure Access data is only through the Secure Lab. Users must conduct all their analysis, and produce their research outputs (such as papers, presentations etc.) within their dedicated project area. Secure Access data will not be released to be used in any other environment under any circumstance.

Users must maintain familiarity with SDC. The two guidelines above offer detail on checking, but users will be given mandatory Safe Researcher Training that covers most of the key concerns around routine analytic outputs. A detailed **User Guide** is also available that Secure Lab users should read and digest. (This guide will also be of value to users of Special Licence data.)

Outputs will be returned to users following a full SDC examination by two trained members of Service staff. It is the policy of the Service to only release 'final results' which are those considered ready for publication or a formal presentation. This avoids multiple attempts at clearing outputs, and is achievable because users working on the same project in the Secure Lab can share their intermediate findings through shared project folders.

Where users are unsure about SDC when they produce outputs, we recommend that they speak to an Service support officer as soon as possible, and certainly before any outputs are submitted for checking. This will avoid disappointment if a user writes an entire paper within the secure environment, only to find that it is not released to them due to SDC problems.

5. Citing data and reporting publications

As outline in the End User licence, all users of data are required to use the DOI of the dataset(s) in any publications or presentations arising from their research to the Service. It is also good practice to inform the Service of any publications at the time of publication and to provide the full citation and DOI. These secondary publications can then to be added to the data's catalogue record to provide useful information for new users.

Owners of data may reserve the right to ask to see or vet drafts of publications based on those data prior to publication. This will be noted in the Access conditions and users notified during the application process.

Users of the Secure Lab are not permitted to publish outputs unless they have been checked and released to them as outlined in Section 4.2.2. Additionally, some data owners may wish to vet publications, and users will be notified in advance.

6. When research is complete

It is recommended that users always retain a well-documented copy of any data preparation or analytic code used to prepare a paper or report.

When a project has been completed, users should remove all copies of the data, including derived datasets, back-ups, paper copies, portable copies (including CDs), and all electronic copies from every device , and any Server, used.

It is essential that all copies of Special Licence data held by users are destroyed and the Service is notified via the completion of a Data Destruction form.

Users of Secure Access data must ensure that their code files are stored in the 'Syntax Folder', which will be retained at the end of the project's life. All other files will be deleted at the end of the project. Code files can be removed from the Secure Lab subject to clearance checks by Service staff.

6.1. Guidelines on data destruction

The following guidelines for destroying data must be adhered to:

- data must be deleted from the system on which it has been stored using a secure erasure programme, such as Eraser (http://www.heidi.ie/eraser/index.php) or similar - which repeatedly overwrites files a number of times, until such time as the original data could not be retrieved forensically;
- the recycle/trash bin must be emptied, preferably to be immediately followed by running a secure erasure programme; portable media holding any data must be returned to the Service or destroyed and disposed of in a secure manner;
- backup tapes must either be completely overwritten and degaussed (demagnetised) before being re-used or disposed of;
- paper copies must be destroyed by shredding, preferably using a cross-cut shredder;
- before the PC, laptop or other device used for data storage leaves the possession of the organisation or individual (for destruction or second-hand sale, etc.), the hard disk must be completely erased using a secure erasure programme;
- destruction of Special Licence data must be confirmed to the Service by the user. A Data Destruction form will be sent to the user one month before the project expires.

7. Organisational responsibilities

UK Institutes of Higher or Further Education (HE/FE) are bound by JANET policies (https://community.ja.net/library/janet-policies/security-policy), including the JANET Security Policy that places responsibilities on every person and organisation involved in the use or operation of JANET to protect the network against security breaches.

Most organisations have their own Information Security policies, which should be consulted before requesting access to Special Licence data. If there is any doubt that your organisation has lower standards than those suggested here, you should check with your IT Services department.

There is a requirement that all central government departments must meet internationally recognised information and security management standards (e.g. ISO/IEC 17799 and/or ISO 27001) for their systems. Local authorities are also obliged to comply with the ISO 27001 security standard as part of their Implementing Electronic Government (IEG) requirements.

7.1. Special Licence and Secure Lab data

For a user accessing Special Licence data, it is the user's responsibility to ensure that they can store and access the data in a suitably secure physical and electronic environment.

Users of the Secure Lab must provide a Secure Access User Agreement which should be signed by an *authorised signatory* of the user's organisation (usually the Registrar or delegate, but not a Head of Department/Centre).

Users of Secure Lab undertake to allow the data owner access to the premises where the data are stored

and/or accessed for the purpose of conducting an audit, without notice and at any reasonable time. (Also see Section 3.5).

Access to Special Licence and Secure Access data may require the user to provide the contact details of a senior member of staff at their organisation who can vouch for their suitability for access to the data. The Service and ONS reserve the right to contact the senior member of staff to ask for a reference.

8. Non-compliance procedures

The user is required to report promptly any non-compliance with any of the terms of the EUL, Special Licence or Secure Lab rules (this includes any non-compliance by someone else that the user becomes aware of). Failure to disclose an act of non-compliance is a non-compliance with the licence.

Non-compliance with the terms of the EUL, including any special conditions, may result in the following actions:

- immediate termination of access to all services provided by the Service and the UK Data Service either permanently or temporarily;
- legal action being taken against the individual who has not complied with the terms of the EUL;
- withdrawal of access to all Service and UK Data Service services either permanently or temporarily to the user's organisation.

Additionally, any non-compliance with the terms of access for Special Licence or secure data:

- will result in the immediate termination of the user's access to the data and the termination of the licence; depending upon the seriousness of the non-compliance, the termination of access may be permanent;
- may result in sanctions being sought against the user by the data owner;
- will, for ONS Secure Access data under the Statistics and Registration Services Act 2007 and the 2018 Digital Economy Act, incur penalties as specified the Acts, which may include a fine and/or imprisonment;
- for Secure Access data, penalties could also include individual or organisational sanctions including withdrawal of UKRI funding and organisational suspension from all UKRI data services.

Users will be provided with detailed guidance on non-compliance and penalties when undertaking the Secure Lab training.

9. Help and feedback

This guide will be regularly updated. For further advice on any of the issues raised, or to provide suggestions or comments, contact the UK Data Service Helpdesk via our 'Get-in-touch' web page: http://ukdataservice.ac.uk/help/get-in-touch.aspx.