UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS
Expert Meeting on Statistical Data Confidentiality
15-17 October 2025, Barcelona

In defence of Scientific Use Files

Cristina Magder, UK Data Service, UK Data Archive, UK
Felix Ritchie, University of the West of England, Bristol, UK
Aida Sanchez-Galvez, University College London, UK
Richard Welpton, UKRI: Economic and Social Research Council, UK

Corresponding author: felix.ritchie@uwe.ac.uk

Abstract

The Trusted Research Environment (TRE) has been the great success story of data access this century. By providing highly secure yet flexible access, the TRE has enabled research use of the most sensitive data.

However, it is the Scientific Use File (SUF) that remains the workhorse of academic research. SUFs are files made available under licence to authorised users, to hold and analyse on their own organisational machines. These are valuable assets to researchers who do not need the detail in TRE data; moreover, they are accessible to users who would not normally be granted access to TREs, such as undergraduate students or non-academic researchers. For organisations like the UK Data Service, which provide research datasets for different uses, the volume of use of SUFs far exceeds that of the TRE.

In recent years there have been concerns about the future of the SUF. A methodological review of confidentiality protection suggested that new technologies and methods can reverse engineer any de-identification techniques. An international conference on microdata access argued that SUFs have a limited future. It has been suggested that, if de-identification techniques are no longer robust, then all personal research data should only be available through TREs.

This would post significant challenges to the research and data services communities: increasing administration, increased costs, with access to fewer users. In addition, the logic of this argument is not just that SUFs are untenable, but that all 'anonymised' data is ultimately identifiable.

This position holds only as long as the argument is limited to the risks to personal privacy. Less often considered, but equally important, is the offsetting benefit to the public of making such data available. Without considering both sides of the equation, arguments about confidentiality risk have little validity. Moreover, there are resource implications for both users and data services: the cost models for SecUFs and SUFs are quite different for all parties.

This paper re-examines this balance between risk and utility, focusing particularly on unrecorded benefits. We note that the public is generally much more supportive of making data available and less risk-averse than data professionals. However, there is a significant gap in evidence: most of the information about public attitudes comes from TRE research and public engagement on SUF research is needed in this area. .

1 Introduction

The Fives Safes model (Desai, Ritchie, Welpton 2015; Green and Ritchie, 2023) is a set of data principles that enable data services to provide safe access to research data. This framework is routinely used to describe how to safely provision access to sensitive data in Secure Use Files (SecUFs, also known as controlled or restricted data). These 'pseudonymised' data (names and direct identifiers removed) provide highly valuable research data, but the minimal extent of de-identification applied to the data leave data subjects vulnerable to reidentification, if the data were to be widely shared. Hence, such data are typically accessed in a controlled manner via Trusted Research Environments (TREs, also known as Secure Enclaves, Secure Data Environments, Research Data Centres). The use of TREs, underpinned by the Five Safes model, has been the great success story of research data access this century, supporting access to the most sensitive data that would otherwise be unavailable for research use.

Yet the advantage of the Fives Safes model, compared to other data sharing governance frameworks, is that it can be used to describe any type of data access route; including datasets made available outside TREs such as Scientific Use Files (SUF, also known as safeguarded data) and Public Use Files (PUF, also known as open data). Dissemination of data as SUF and PUF accounts for the vast majority of granted data access. The data themselves are typically accessed via download from repositories such as the UK Data Service.

And it is the SUF that remains the workhorse of academic research. SUFs are files that contain individual-level data that have been subject to much more extensive de-identification methods (compared to SecUFs) to reduce data disclosivity. Outliers may be removed, values may be truncated, and geographic information may be limited. Precise values might also be replaced with percentiles, group averages or banding. In some cases, records may be swapped. Some data producers test the data to ensure that a sufficient k-anonymity test is met. But there may still be enough information, with a small remote probability, to identify a data subject. Consequently, the data cannot be made available through unrestricted release (as in PUF or open data). Instead, the data are made available as an SUF, under licence (with multiple and robust conditions of use) to authorised users, to hold and analyse on their own servers.

In the UK, the UK Data Service is the largest of a handful of European data infrastructures that provide access to data across the full spectrum (PUF, SUF and SecUF). The volume of use of SUFs is roughly twenty times that of TRE SecUF files. SUFs are valuable assets to researchers who do not need the detail of the sensitive data only available in SecUFs. SUFs are generally very accessible to researchers, whereas access to sensitive data in TREs requires, understandably, much more scrutiny and bureaucracy. Some researchers, such as students, are barred from applying to TREs on the basis of insufficient experience. The time taken to apply for access to TREs (which in some cases can take years) can prevent even experienced researchers from undertaking research in respond to national priorities.

However, there have been concerns about the future of the SUF. A methodological review of confidentiality protection (Ritche and Smith, 2019) suggested that new technologies and methods (such as large-scale computing and AI, combined with social media availability) can reverse engineer any de-identification techniques applied to whole-population samples (e.g. administrative data) by the data producers. An international conference on microdata access (Green et al, 2021) suggested that SUFs have a limited future; one head of a major data provider made similar comments. It has been suggested that, if de-identification techniques are no longer robust, then all personal research data should only be available through TREs.

This would pose significant operational and administrative challenges to the research and data services communities. It would also limit research use to a privileged few. A generation of researchers from undergraduates upwards would struggle to develop their applied statistical/research skills using easily available heavily de-identified data, with consequences for the research workforce and national research capability. Finally, the logic of this argument is not just that SUFs are untenable, but that all 'anonymised' data is ultimately either identifiable or of such low specificity as to lose any research value.

This position rests upon assumptions about technological development; but its fundamental weakness is that it is framed as if personal privacy is the only factor of interest. Less often considered, but equally important, is the offsetting benefit to the public of making such data available. A separate legal argument can be explored concerning whether data made available via SUF are 'personal data' or not. This is a tangential aspect, but the purpose of our paper is to re-examine the balance between risk and utility. Without considering both sides of the equation, arguments about confidentiality risk have little validity.

2 File types and release mechanisms

Because multiple terms may be used to describe the same object, we define the following file types and some example synonyms, specifically those from the Turing Institute, UK Data Service (UKDS) and the UCL Centre for Longitudinal Studies (CLS). While there are subdivisions of these, this tripartite split is sufficient for the discussion in this paper.

Name, acronym (as used in this paper)	Description	Synonyms
Public use file PUF	Open access, made available with no significant restrictions; data subjects cannot reasonably be re-identified	Open (UKDS) Tier 0 (Turing)
Scientific use file SUF	Available for researchers to acquire and keep on their machines subject to agreed conditions described in a licence; some restrictions around access and use; data confidentiality risk assumed to be a low but non-negligible risk of re-identification	Safeguarded (UKDS) Tier 1 (CLS) Tier 2 (Turing)
Secure use file SecUF	Researchers need to access data in a TRE. No options to upload or download data; data subjects are assumed to be easily identifiable.	Controlled (UKDS) Tier 2 (CLS) Tier 3 (Turing)

Two other common terms used are 'identified' (with specific identifiers such as name or social security numbers included) and 'pseudonymised' (removal of direct identifiers). Identified data is rarely used in research as the identifiers have little research value; if they are used, they are used in secure facilities. Secure facilities are most likely to hold identifiable but not identified (i.e. pseudonymised) data. Pseudonymised data, meanwhile, might include both SUFs and SecUFs.

3 The data access spectrum

A useful way to view differences between file types is the 'data access spectrum'. Green and Ritchie (2016) illustrate for a hypothetical dataset how the same information may be made available for researchers in multiple ways, illustrated in Figure 1 below:

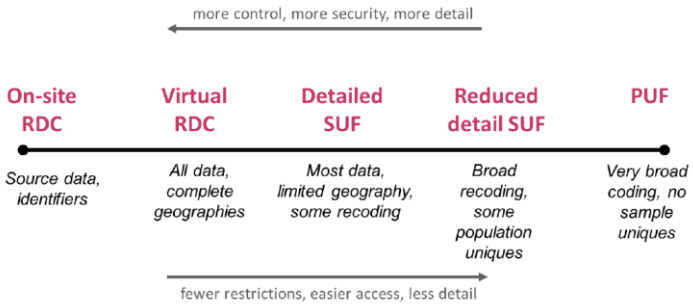

Figure 1 Data access spectrum (Green and Ritchie, 2016)

The value of this comes in the context of the Five Safes (Desai et al, 2015). This is a common framework for describing data access solutions which considers five dimensions of risk control: projects, people, setting, outputs and data. The Five Safes can be combined with the Data Access Spectrum by contrasting the data controls (where a dataset is on the spectrum) with the procedures/technical controls (the other four safes). For example, Ritche and Kendal (2024) illustrate with reference to the level of controls in that dimension (controlled>checked>trusted>no effective controls), compared to the level of data detail reduction (none=source data, complete=safe for unrestricted release); see Figure 2.

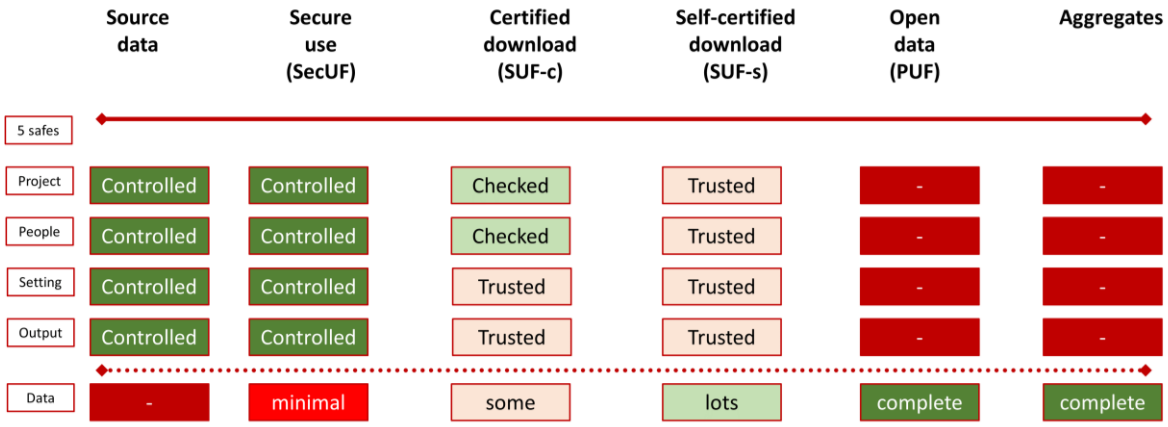

Figure 2 Five Safes controls applied to the data access spectrum (Ritchie and Kendal, 2024)

The value of this perspective is that it makes clear that there are choices to be made, and that multiple solutions can meet the goal of "safe use". Each solution embodies different perceptions of what combinations of controls works to protect privacy, and what the implications are for users. See the Appendix for a specific example from the UK Data Service

3.1 File use

Figure 3 shows the 'usage' rates for different file types in the UK Data Service from April 2019 to March 2025:

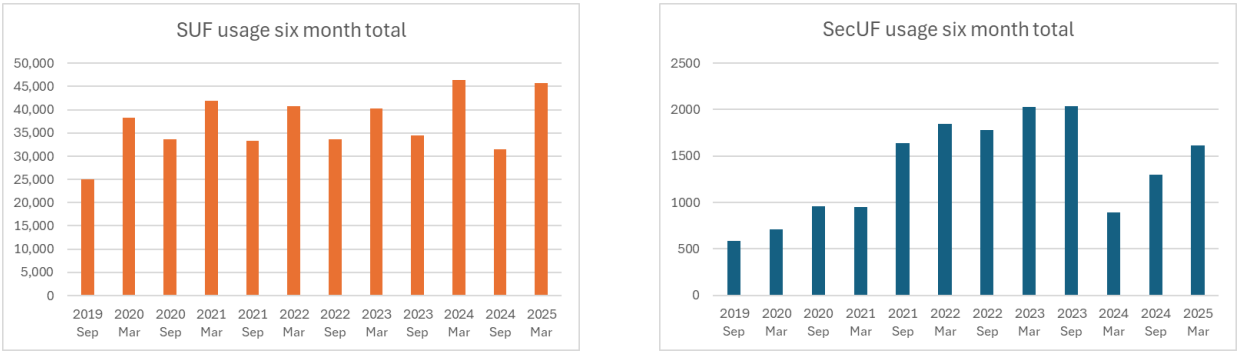

Figure 3 Usage data from UKDS for SecUFs and SUFs

The figures show download from the repository for a user, either for offline use (for SUFs) or into the research space (for SecUFs). Hence the figures do not show how much each dataset is used. SUFs are more likely to be downloaded and not used, following examination by the potential user. So these figures are a rough estimate of demand for the datasets, but cannot exactly be compared.

Despite these caveats, it is clear that SUF usage is an order of magnitude greater than SecUF. SUFs are also used by a wider range of users. SecUF files are limited to those who can demonstrate appropriate qualifications, and who can afford to wait for sometimes lengthy application processes. In contrast, SUFs are available for non-commercial and teaching use to undergraduates, independent researchers, charities and think-tanks, journalists, researchers in other countries, and so on. SUFs might also be available for commercial use.

3.2 Examples

To further investigate proportionality in practice, we examined usage for two UK data series disseminated by the UK Data Service at multiple access levels, one longitudinal and one cross-section. In each case, it is the same underlying data resource, but the versions offered scale with sensitivity and use case. Because UKDS collects information on users when they register for the download licence, it is possible to analyse the type of uses, as well as to filter out multiple downloads to the same user.

3.2.1 Example of secondary usage for a longitudinal data series

The longitudinal study is disseminated at three levels: SUF, SUF+ (still downloadable, but with additional restrictions on access and use), and SecUF. Between 2020 and 2025, this collection had 17,295 unique SUF users (90%), 1,738 SUF+ (9%), and 216 SecUF users (~1%). The distribution shows that most researchers can proceed with their project using effectively anonymised data, with escalation only where finer details or higher sensitivity are justified. SUF use is broad and routine, with higher education dominating (~78%), followed by additional uptake from further education, government, and the third sector. A small number of commercially affiliated researchers also access the data for non-commercial purposes. The SUFx version serves a smaller, more specialised audience (1,738 users, again ~80% from HE).

The SecUF contains the most sensitive and granular elements and is only available for approved projects in provider-managed secure environments, with mandatory training and output checking. This route enables work that truly needs the highest granularity such as modelling rare outcomes under tight governance.

3.2.2 Example of secondary usage for a cross-sectional data series

The second example is released as PUF, SUF, and SecUF. For the PUF only downloads can be counted, as no information is recorded on the user types or usage type. From 2020 to 2025, the PUFs were downloaded 1,583 times. Downloads peak in 2021 (326) and then stabilise at ~300 per year. This tier supports discovery and teaching prep but, by design, lacks the micro-level detail needed for more in depth analysis.

The SUF drives the bulk of research use: 6,568 unique users across 5 years. Nearly all use is non-commercial or educational, and the user community is overwhelmingly comprised of academics and students, accounting for over 90% of all users. There is also meaningful uptake from central/local government and the third sector.

During the same period, from 2020 to 2025, 346 unique accredited users accessed the SecUF, with use similar to that of the SUF: ted in HE/FE, followed by third sector users.

Taken together, these two examples further show the broad value realised safely by using a tiered data access spectrum and moving away from dichotomies such as open or closed.

4 Assessment of Scientific Use Files and Secure User Files

4.1 Context

In recent years, some data producers, concerned with the potential risk of reidentification, have determined that data once previously available as an SUF, should now be accessed only in a TRE as a SecUF. In cases where this has happened, research access has dramatically fallen. For those data, less research is undertaken with the

data. Is this a price worth paying if we believe that research benefits society? There is a risk that data access will be polarised between secure and open, with no intermediate options between the two. This provides the context that underlines our contention that SUFs remain a valid – and vital - approach to balancing data confidentiality with utility.

4.2 Benefits of maintaining the data access spectrum

4.2.1 Value

The purpose of creating research datasets is to support research for public benefit. That benefit comes from (a) providing researchers with sufficient detail, in reasonable time, to undertake analysis that is valuable (b) ensuring that the data is available to those who can use it for public benefit.

The value of social science research, though not the topic of this paper, can be assessed by examining a range of impact case studies recording by data service infrastructures such as UK Data Service. Alves et al (2021) demonstrate that trying to put a financial value on this is largely impossible or meaningless. However, noting the usage figures in the previous section, even if a fraction of access requests result in a scientific publication, considerable scientific endeavours are progressed from the availability of data through the SUF access route.

Public benefit is also derived from the creation of a skilled data analytics workforce. Particularly in social science, researchers from an early career stage (e.g. undergraduate) begin applying statistical techniques learned in the classroom to real world data, with the many advantages this confers later on.

Public benefit comes from having data available to make evidence-based policy; but often forgotten is the counterpart. There is a cost to society of *not* having data available, and therefore policymaking being based on hunches or predetermined political positions. We are not aware of any studies that try to put a value on either capacity building, or on foregone benefits from lack of data.

Public benefit is of course also created when researchers access SecUF data in TREs, where they gain access to sensitive data which otherwise could not be used for research, to the benefit of our understanding of society; but as the usage statistics above show, fewer researchers access data this way. TREs as technical solutions to data access can be limited: for example, software licences may restrict the number of researchers who can log in to use SecUF data in a TRE at any one time. The data governance requirements that must be met and processed to enable access to SecUF data in a TRE in the first place restrict the flow of researchers using the data. All this means that while the *value* of TRE-enabled research maybe higher than for SUFs, the *incidence* of it is much lower, and the *opportunity* for capacity building is much more limited.

4.2.2 Costs

As with any statistical data release ecosystem, costs must be considered when provisioning access to data. Here, the SUF has clear advantages. Creating an SUF incurs a one-off fixed cost, and preparing the data for research (e.g. perturbing the data to protect confidentiality) can be done as part of other quality control aspects and documentation creation. Once created, the dataset is made available for download, subject to registration and agreement to licence conditions, at no cost to the researcher.

A SecUF needs less preparation. For example, removing direct identifiers such as names, addresses etc. is usually only the necessary requirement (and often these are not useful for research anyway). The rest of the data are usually left intact. But provisioning access through a TRE can be costly: the technical set-up, ongoing information governance accreditation, providing staff to support researchers, software licences etc. all mean that the marginal cost of access is relatively high compared to the SUF case where a researcher can simply download the data to their own computer.

Therefore, SUFs, while they incur a relative high fixed cost of preparation compared to a SecUF (which have smaller preparation costs because less perturbation is required), can be disseminated for a small or negligible marginal cost. The researcher does the work of downloading the data to their own machine and managing it.

4.3 The risks of data dissemination

The arguments against releasing files for researchers via SUF have two bases: that the controls are inadequate, and that the data is insufficiently protected against bad actors.

4.3.1 Inadequate controls

The controls on projects, people and settings largely rely upon the goodwill of researchers. Researchers downloading SUFs are typically asked to identify the purpose for which they intend to use the data and to identify which organisation they work for. They are usually asked to make a commitment to holding the data in a secure location (such as a password-protected device or server) and not copying onto portable media). Often a data use licence is agreed to, which could lead to legal repercussions if broken, although the enforceability of these may be questionable. Researchers accessing SUF data may sometimes be given guidance of the disclosure risks in outputs, but this is rare (Derrick et al 2024), and there is no evidence that researchers read the guidance.

But likewise, access to TREs, with their stringent criteria, and often a requirement to attend training, are not infallible. It is perfectly possible that a researcher undertakes all that is requested of them with respect to agreeing to terms of use and training; but then writes down results from a screen and publishes them, going unnoticed by TRE staff. The *degree* of control changes but not the ultimate reliance on researcher training.

There is limited opportunity to detect if researchers choose to ignore these restrictions in the SUF download or TRE access. It is unlikely that a dataset being used for a slightly different purpose than the one proposed would be spotted in either scenario; data services retain the right, but do not usually carry out random checks on the computers belonging to SUF researchers.

Nevertheless, there is some evidence, particularly from psychological studies, that asking people to make these or similar commitments does affect behaviour positively. A problem is that there is almost no evidence relating specifically to the use of SUFs by researchers. We believe that researchers are well-intentioned but may sometimes be careless and self-centred (Green et al, 2017). On the other hand, many researchers, valuing the ease of access offered by SUFs, have anecdotally told the authors of their commitment to abiding by conditions of access, when they understand the consequence of such data no longer being available. But the evidence for our assertions is largely anecdotal or informal; this is a significant gap in our evidence base.

For these reasons, data detail in SUFs is significantly reduced (as illustrated in Figure 2) so that, if a researcher is malicious or just careless, the likely loss of confidentiality is small. It is not reduced as much as it is in PUFs, because we do think the measures are partially effective.

4.3.2 Data detail

Ritchie and Smith (2019) used a simple model to show the inevitability of being able to reverse-engineer any data to identify the data subjects. Computing power and auxiliary sources of information, such as social media, are effectively unlimited in the long term. Meanwhile, the ability to reduce detail in the data is construed by the dataset itself. Ultimately, the former will overtake the latter, unless the data detail is reduced so much that it effectively becomes aggregate data – k-anonymity on every combination of variables.

There are useful things that can be said about the practical risk involved in any dataset. Assessment of re-identification risk is a mature field of research. The increasing availability of external datasets makes some of the less realistic models of twenty years ago more sensible now. The disclosure risks in increasingly complex datasets (particularly the growth in longitudinal studies) are significant. But neither of these changes the argument that we can, meaningfully, make a judgement about the re-identification risk.

Yet we contend that when actual data disclosures are reported, there can be misunderstanding about the nature of the data released. Often data that fall into the wrong hands or used maliciously are described as anonymised, but on closer inspection these were not SUFs, as very little, sometimes if any, de-identification techniques and data use controls have been applied to the data. In a recent case, the data were identifiable: they contained full names. Describing these as anonymised data is false; and similar cases that are used to describe the fallibility of SUFs could irreparably damage research if these examples are used to stop disseminating data as SUFs.

On the other hand, if SUFs did not exist, and researchers only accessed SecUF data in TREs, would they be accessing more detail than required for their research? The usage figures illustrated above provide that SUFs are adequate to the needs of many researchers. Remove the availability of these data, force researchers to access SecUF in TREs, and they may access more information than intended otherwise; going against the spirit, if not the letter, of data protection legislation or at least, best practice.

4.4 Costs versus benefits: the need for balance

The counter-argument to the previous section is that it is a one-sided argument. If the sole objective is to prevent confidentiality breaches, then not releasing data is the only guaranteed solution. This ignores the public interest arguments for enabling access to data for research. Ultimately, there are costs and benefits to the public of both releasing and not releasing data for research, and SUFs and SecUFs have *different* cost-benefit profiles, not unambiguously better or worse ones.

5 Public acceptability

Members of the public, when asked, tend to be much more prepared to accept risks to their data to create public benefit than data professionals are prepared to accept. Public engagement research repeatedly supports this. Given that public benefit is the goal, these public views are crucial for considering the risk/value argument.

In recent years, UK funding sources' condition of funding to include public engagement has generated much useful knowledge. Unfortunately, the focus on support for TREs means that much of the public engagement asks about TRE use (Kashef et al., 2022). As TREs have an easy story to tell about safe data management, the lessons from such public engagement research cannot directly translate to SUFs. And it may be more difficult to discuss concepts of anonymisation related to SUF data preparation to the public relative to making security assurances about TREs. This is a significant gap in knowledge.

It is also well-known that framing affects perceptions of any decisions. The focus on risk in discussions around distributed data such as SUFs, rather than on the value of such data, pre-disposes participants in those discussions to start from a default-closed perspective i.e, nothing will be done unless it can be proved safe. This alone leads to under-provision of data resources for society's goals.

6 Conclusion

For many decades, SUFs have enabled much research to be undertaken for the benefit of society; and played a major role in training our research workforce and analytical capability. However, the evolution of technology and increase availability of data in the public domain represent increased risks to the confidentiality of the data and identifiability of the data subjects; and therefore, challenges the broad data access spectrum. On balance, maintaining SUFs alongside TREs maximises public value while managing risk proportionately.

What is needed next is a conversation with the public about the effectiveness of de-identification and anonymisation; and risks they are willing to take in the interests of scientific research to continue, which they and future generations will benefit from. Further work to understand the exact risks *and* benefits of disseminating SUF files should also be undertaken.

References

K. Alves, F. Tava, D. Whittard, E. Green, M. Beata Kreft, and F. Ritchie. Process and economic evaluation of the ODI R&D programme: Final report, 2021. Open Data Institute, London. https://uwe-repository.worktribe.com/output/7294686

Derrick B., Green E., Ritchie F., Smith J., White P. (2024) The inadvertently revealing statistic: A systemic gap in statistical training?, Significance, 21:1 pp24–27 https://doi.org/10.1093/jrssig/qmae009

Desai T., Ritchie F., and Welpton R. (2016) The Five Safes: designing data access for research. Working papers in Economics 1601, University of the West of England, Bristol.https://uwe-repository.worktribe.com/output/914745

Green E. and Ritchie F. (2016). Data Access Project Final Report. Australian Department of Social Services https://uwe-repository.worktribe.com/output/908255

Green, E. and Ritchie F. (2023) "The Present and Future of the Five Safes Framework". Journal of Privacy and Confidentiality 13 (2). https://doi.org/10.29012/jpc.831.

Green E., Ritchie F., Newman J., and Parker T.. Lessons learned in training ''safe users" of confidential data, 2017. UNECE Work Session on Statistical Data Confidentiality 2017.

Kashef, S., & Cowan, M. (2024). Exploring UK-wide public perceptions of the 'public good' use of data for research and statistics. International Journal of Population Data Science, 9(5). https://doi.org/10.23889/ijpds.v9i5.2735

Ritchie F. and Kendal C. (2024) Future Data Services Briefing note 6: The data access spectrum. https://uwe-repository.worktribe.com/output/13507443

Ritchie F. and Smith J. (2018) Confidentiality and Linked Data. In: National Statistician's Quality Review, ONS, Newport, December. https://uwe-repository.worktribe.com/output/856040

UNECE

Appendix: using the five safes to assess file types in the UK Data Service

The UK Data Service has a formal matrix for identifying how different file types are mapped to the Five Safes controls. With the exception of PUFs (for which they are not relevant) these conditions are specified in the data access agreements with researchers.

Five Safe	Open (OGL/OPL/CC BY) (PUF)	Safeguarded (EULA*) (SUF)	Safeguarded Special Licence (EULA + SLUA**) (SUFx)	Controlled (EULA + SAUA**) (SecUF)
Safe Data	Not Personal Data or Information with no real disclosure risk e.g. aggregate data, samples etc. or Personal Data/Information with consent to share as collected under Open access.	Not Personal Data or Personal Information; effectively anonymised with sufficiently remote reidentification risk.	Not Personal Data or Personal Information; effectively anonymised with sufficiently remote reidentification risk, with more granular detail (e.g. lower-level geographies such as Local Authority).	Personal Data or Personal Information where direct identifiers have been removed or otherwise highly sensitive data commercially or otherwise.
Safe People	No restrictions; anyone may access the data.	Registered user; awareness of ethical and legal data handling expected; training and guidance available; specific user type restrictions may apply.	Registered user; awareness of ethical and legal data handling expected; training and guidance available; specific user type restrictions may apply.	Registered and accredited users; awareness of ethical and legal data handling demonstrated; training mandatory; other specific user type restrictions may apply.
Safe Projects	No project registration or approval required; reuse permitted under licence T&Cs (e.g. attribution, no commercial use, no derivative works etc. as applicable).	Project registration required; use must be specific and time-limited; project application and formal approval may apply.	Project registration and application required; use must be specific, time-limited, and formally approved by data owner/their nominee.	Project registration and application required; use must be specific, time-limited, with clear public benefit and formally approved by data owner/their nominee.
Safe Settings	No restrictions; data may be used in any environment.	User-managed setting; e.g. safe device/end point with data only accessible to the registered users and securely deleted after project end, as per licence terms.	User-managed institutional setting; access via declared institutional device only (no personal devices); additional agreement and permission if working from home; data destruction form required.	Data Service Provider–managed secure setting (e.g. SecureLab, SafePod, IDAN); access only via approved institutional devices to a monitored and locked-down environment; additional agreement and permission if working from home.
Safe Outputs	No restrictions or requirements.	User responsible for ensuring outputs are non-disclosive (e.g. min threshold 3 for primary, 10 for secondary, region as geography etc.); additional outputs conditions may apply.	User responsible for ensuring outputs are non-disclosive (e.g. min threshold 3 for primary, 10 for secondary, region as geography etc.); additional outputs conditions may apply.	All outputs are checked and approved by the Data Service Provider before release; additional outputs conditions may apply.

* https://ukdataservice.ac.uk/app/uploads/cd137-enduserlicence.pdf
** https://ukdataservice.ac.uk/app/uploads/special_licence_application_bundle.zip
*** https://ukdataservice.ac.uk/app/uploads/cd140-secureaccessagreement.pdf