



*The Creative Commons-Attribution-Noncommercial License 4.0 International applies to all works published by IASSIST Quarterly. Authors will retain copyright of the work and full publishing rights.*

## Developing canonical ‘safe researcher’ training materials for trusted research environments

Deborah Wiltshire <sup>1</sup>

### Abstract

Social science and humanities research infrastructures allow the sharing and safe use of confidential, sensitive data for research via physical safe havens. In recent years there has been a shift towards virtual data enclaves or Remote Desktop systems that offer fewer physical controls. These controls need to be replaced with other safeguards, including mandatory ‘Safe Researcher’ training. This training aims to ensure that researchers are equipped with the knowledge required to use secure data safely. Developing training is resource intensive so canonical training materials are an economical approach to providing standardized, high-quality training.

The Social Sciences and Humanities Open Cloud project deliverable ‘Training materials of workshop for secure data facility professionals’ had two objectives. The first was the development of a set of canonical training materials that Trusted Research Environments (TREs) could use as a framework on which to build their own training course. The second objective was to hold a virtual workshop where the training materials could be demonstrated to a credible audience to gather feedback to inform the future development of the materials.

We have now developed the canonical materials, building on the wealth of expertise and experience of UK-based TREs. These training materials were then demonstrated at a virtual, two-hour Stakeholder Workshop that we organized in September 2021. Following our demonstration of the materials, we facilitated small group discussions to gather vital feedback. The discussion groups formed a consensus that the materials were both comprehensive and clearly structured and would be a valuable resource to the TRE community.

### Keywords

Safe researcher training; canonical training materials, sensitive data, Trusted Research Environments

### Introduction

Social science and humanities research infrastructures provide a variety of resources and services that researchers can use and benefit from. They provide infrastructures that allow the sharing and safe use of confidential, sensitive data for research. These infrastructures are often in the form of Trusted Research Environments (TREs) that provide safe data access and use by creating highly secure digital

environments. These digital environments are augmented by non-technical controls, and TREs often choose to adopt the Five Safes Framework as the basis of their security model. The Five Safes Framework details five principles – safe projects, safe settings, safe data, safe people, and safe outputs – which can be successfully balanced to ensure the safe use of sensitive data (Desai, Ritchie, and Welpton, 2016; Woollard et al., 2021).

Originally access to these confidential data was only via Safe Havens or Safe Rooms - physical, secure rooms often located within a Data Archive or Secure Data Center<sup>2</sup>. These rooms are specially designed to ensure that strict physical controls are in place. Such controls include lockable rooms with restricted access, the barring of personal items such as electronic devices in the Safe Room, and the prohibiting of taking handwritten notes whilst working in the room. In recent years there has been a shift towards virtual data enclaves or Remote Desktop systems via TREs. These systems offer more flexibility as researchers can access the secure digital environments from their own institutional offices via a secure VPN connection, so they are highly popular. But they also introduce the potential for greater risk as they offer fewer physical controls. The controls lost need to be replaced with other safeguards in line with the Five Safes Framework. These safeguards include legal measures in the form of licenses, data use agreements, and contracts – mandatory Safe Researcher training for researchers prior to data access being granted is often included as well.

It is widely recognized within the TRE community that the researcher is an integral part of any security model and should contribute to mitigating the risks associated with disclosive and sensitive data (Lambert 1993; Wiltshire 2022; Bishop et al, 2022). Therefore, the function of Safe Researcher training is to ensure that researchers have the knowledge and the appropriate attitude to avoid mistakes or poor practices that might otherwise lead to a data confidentiality breach (Desai, Ritchie and Welpton, 2016). Many TREs in the UK have mandated Safe Researcher training as part of their security models for many years, and whilst some in the wider European TRE community also offer training, it is not yet adopted as widely. This training can cover a range of different topics depending on the specific requirements of the TRE but will typically include information on the legislative frameworks or constraints, service-specific information, and statistical disclosure. This ensures that researchers are equipped with the knowledge required to use sensitive data safely. The experience of TREs across the UK and beyond shows that such training can have a significant positive impact on protecting the confidentiality of the data (Bishop et al., 2022). One example is the UK Safe Researcher Training which is a half-day training course complete with assessment which forms part of the Accredited Researcher Scheme<sup>3</sup> overseen by the Office for National Statistics. All researchers who wish to access sensitive data made available under the Digital Economy Act must undergo this training. The training covers topics such as data security and the researchers' responsibilities and statistical disclosure control, as well as service-specific information to help researchers use the service more effectively and efficiently. The training can be developed by any TRE who make these data available, and this consortium-based approach enables researchers to train with one service but carry their 'trained' status to another service within the group.

As TRE community increasingly switches to remote access, rather than on-site access, discussions have turned to introducing training into our security models. With the additional driver of increasing remote

access connections that allow access to sensitive data across international borders, ensuring that we have consistent and comparative security models across the different TREs is becoming a key priority. Having some commonalities in the training that TREs across the world offer can only be a positive addition to ongoing efforts to open up sensitive data access across international borders. The desire to implement training is not without challenges. For TREs outside of this UK-based group, implementing such training would involve starting from scratch and this is potentially a prohibitive barrier for many. Developing any training course is resource intensive, and for smaller TREs with limited resources, designing and developing a Safe Researcher training course from scratch is a burden not easily overcome. Looking for ways to overcome this barrier to implementing Safe Researcher training was the focus of one of the working groups of the Social Sciences and Humanities Open Cloud (SSHOC) Project, a large project that aimed to expand access to sensitive data in Europe.

### **The Social Sciences and Humanities Open Cloud Project**

The [SSHOC](#) project<sup>4</sup> was an EU funded project that brought together 47 organisations from across Europe and from across disciplinary boundaries to advance and to contribute to the work of the [European Open Science Cloud \(EOSC\)](#)<sup>5</sup>. The SSHOC partners brought a breadth and depth of expertise and experience across the entire data cycle from data collection and curation to data re-use and training.

The SSHOC project ran between January 2019 and April 2022 and focused on transforming the heavily siloed social sciences & humanities data landscape into an integrated, cloud-based network of interconnected data infrastructures. It consisted of 9 work packages encompassing a range of different deliverables and milestones. Work Package 5 focused on innovations in data access and includes deliverables aimed at enhancing and extending the infrastructure for secure remote access to sensitive data. As part of this Work Package, myself and colleagues from TREs across Europe worked to deliver several key deliverables aimed at tackling some of the challenges and barriers to international access. In particular we were keen to find a way to facilitate TREs, especially those with fewer resources, in implementing their own Safe Researcher training via the deliverable D5.20 'Developing canonical training materials'. This deliverable had two objectives: the first was the development of a set of canonical Safe Researcher training materials that any TRE looking to develop Safe Researcher training could use as a framework on which to build their own training course. The second was to hold a virtual workshop to debut the training materials in front of a credible audience of secure data access professionals, trainers, and researchers in order to gather feedback on the materials.

### **The SSHOC Safe Researcher canonical training materials and workshop**

As the lead of the Secure Data Center team at GESIS with many years' experience of delivering Safe Researcher training in the UK, the task of overseeing the development of the SSHOC canonical training materials fell to the author. This was opportune as while mandatory Safe Researcher training is not yet routinely in place at TREs in Germany, one of our priorities is to introduce it at the Secure Data Center. But like many TREs, we are a very small team and outside of the SSHOC project, we would not have had the resources to develop our own training.

When considering how to approach this task, our early decisions were heavily influenced by the

successful development and implementation of the Safe Researcher Training (SRT) scheme in the UK. Under this scheme, a canonical set of training materials were developed by the team at the Office for National Statistics (ONS) and then used by multiple UK-based TREs to deliver their own training courses. I had worked with these training materials and delivered them to the research community over many years and knew that they both covered the topics required by TREs and were well received by participants. Therefore, as a project team we were confident that canonical training materials would be an economical approach to providing standardised, high quality Safe Researcher training programmes across multiple TREs, and that the ONS SRT scheme provided a suitable model for us to adopt.

Safe Researcher training is a very niche area, and to date there are very few practical examples to follow and not a large body of research available to us. Therefore, as a starting point for developing our own canonical training materials, we approached Felix Ritchie, the lead author of the ONS' Safe Researcher Training program to ask if we could use his materials as a basis for our own materials. We made the decision to base our own materials on these resources, because these are tried and tested over many years, and have evolved based on the experiences and feedback of both trainers and participants alike. Whilst there is to date no formal research specifically into the efficacy of these training materials, anecdotally those who work in TREs that implement such training, see a positive difference in key markers such as researcher attitudes towards data governance and the quality and safety of outputs.

In July and August 2021, I developed a set of SSHOC canonical set of materials consisting of a 94 slides PowerPoint presentation. Where appropriate, I added speaker notes that were designed to aid future trainers in delivering the key messages. Following the completion of the SSHOC project, we made the materials publicly and freely available via the SSHOC projects' Zenodo account (Wiltshire, 2021a)<sup>6</sup>.

Whilst the core content of the materials is very similar to the UK materials, there are a couple of key differences. Firstly, the UK materials were developed by the team at the ONS specifically for UK TREs that make data available under the Digital Economy Act. In contrast, with our materials we wanted to make them accessible to a more general audience so that they could be adapted to make a wide range of needs. Secondly, the ONS training materials are designed to be largely a finished course, with other TREs adding a few additional slides to give service-specific information as the only edit. With our materials, our main aim was to produce something that TREs could use as a starting point for developing their own training course. So, the materials provide information on core topics that could easily be adapted to the specific needs of TREs across different countries and potentially different data types. In this sense, the materials are designed not to be a finished product but to provide a more flexible framework upon which TREs can build their own training course.

### The structure of the training materials

The training materials are arranged in six modules, each covering a distinct topic as follows:

- Module 1 - Introduction
- Module 2 - Understanding what impacts data access

- Module 3 - The role of legislation in data access
- Module 4 - The Five Safes framework
- Module 5 – Statistical Disclosure Control
- Module 6 – Service-specific protocols

I arranged the content into this modular format primarily to provide a clear structure and easy adaptability. I designed the module structure so that each module both builds on the previous module and provides the key concepts required to understand the next module. For example, module 2 discusses some of the factors which impact data access including the role of legislation which is discussed in more detail in module 3. Having discussed these factors, the materials then move on to introducing the Five Safes Framework as a way of thinking about and managing data access. This provides a clear, logical path through the materials that makes it easy for researchers to follow. The modular approach should also allow for easy adaptability, allowing TREs to more easily identify and isolate content that they need to adapt.

### Considering course delivery modes

Initially Safe Researcher training courses were run as in-person courses led by one or two trainers. During the recent COVID-19 pandemic, training shifted to virtual delivery with great success and researchers were appreciative of not having to travel. Although the pandemic is over, many TREs have opted to stay with virtual delivery as this is a cheaper and more convenient option for both them and the participants.

A consideration should be given to the course length, especially when thinking about virtual delivery. Experience shows that virtual courses often need to be shorter to avoid 'Zoom fatigue', therefore I developed these materials to function as a taught course that could work equally for in-person or virtual delivery. To aid TRE teams in adapting the materials for virtual delivery, the slides include comments and recommendations on where changes may be appropriate. For example, suggesting content that could be moved from the presentation into a supplementary handout to reduce the overall course length (as seen in figure one). Elsewhere, two different versions of a slide have been included and guidance given on which to choose for specific audiences.

Figure 1 Excerpt from the canonical Safe Researcher training materials

## Legal gateways, useful definitions

Optional: can be included in a course handbook (see notes)

**General Data Protection Regulations (2016)** The GDPR is wide-reaching; main concerns include the rights of data subjects, the duties of data controllers or processors, and the liability or penalties for breach of rights

**Common Law:** The law derived from decisions of courts and case law, rather than Acts of Parliament or other legislation.

**Duty of Confidentiality:** A duty of confidentiality arises when one person discloses information to another in circumstances where it is reasonable to expect that the information will be held in confidence.

**Confidentiality:** Ensuring that information is not made available or disclosed to unauthorised individuals, or organisations.

**Implied consent:** An unwritten 'agreement' between the patient and health and social care professionals that provide their care that allows their data to be shared as long as it is relevant for their care.

**Explicit consent:** A freely given, specific, informed and unambiguous indication of the individual's wishes e.g. regarding data use.

[Welpton, R., Kotrotsios, Y., Parker, S. Data Awareness Training \(securedatagroup.org\)](#)

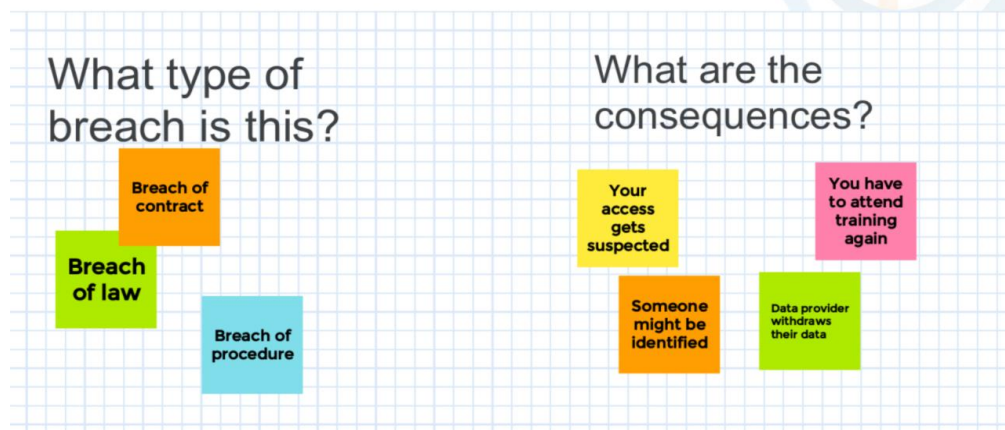
This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782



Another key feature of the training materials are suggestions for group exercises and where such exercises could be included. Group exercises can play a key role in ensuring the participants are engaging with the materials and in helping embed the knowledge using practical exercises. These group exercises work well in an in-person setting, however, can be difficult to facilitate in a virtual setting. During the demonstration of the materials, discussed in more detail in the next section, I demonstrated how interactive whiteboards such as JamBoard can be used during virtual courses to allow participants to collect their ideas and responses in a communal location (figure two)<sup>7</sup>.

Figure 2 Excerpt from the canonical training materials showing an interactive group exercise

## Breach Exercise - JamBoard



## The SSHOC workshop ‘Developing Canonical Training Materials’

Once the development of the materials was complete, the SSHOC project team organised a workshop entitled ‘Developing Canonical Training Materials’ which took place on 21 September 2021. We sent invitations to stakeholders within the SSHOC and TRE communities, and the SSHOC team also advertised the workshop online via the project website<sup>8</sup>. As the key person involved in the development of the materials, I hosted the workshop, which due to the ongoing Covid-19 pandemic, was held virtually via Zoom. The workshop lasted around two hours and designed to be fully interactive with a targeted audience consisting of TRE professionals, experienced trainers, and researchers. In total 22 people from across Europe and the United States attended the event.

The workshop was divided into two parts – during the first part of the workshop I presented the materials, discussing the motivation behind the development of the training materials and their proposed benefit to the TRE community, followed a demonstration of how materials would be delivered. Going through each of the six modules in turn, I talked through both the purpose of the module and the content, pointing out along the way where the content or the delivery could be adapted to suit individual TREs. In the second part of the workshop, I divided the participants into small groups of 4-5 and sent them into breakout rooms for discussions focusing their thoughts about, and recommendations for the training materials.

### Key themes from the discussions

Each group was presented with initial questions aimed to stimulate the discussion and steering it to the specific areas where we were particularly interested in getting feedback. We were particularly interested in gathering their thoughts around the following areas:

1. What do you think worked well?
2. Do you think that the course structure is clear?
3. What do you think about the content? Is it clear? Is it comprehensive?
4. Are there any topics that you feel are missing from the materials? What else would you need if you were delivering these materials?
5. Any other comments?

Participants were given 30 minutes for the discussions to ensure that they had enough time to freely share their thoughts and ideas. All participants were very open to learning about the materials and this made for very lively and engaged conversations. Several key themes emerged in the discussions which are summarised here. This summary includes also any recommendations made and our follow-up thoughts or actions.

### 1. The focus of the materials

The first key theme centered around the focus of the materials. The primary focus of the materials is on quantitative analysis and therefore quantitative research which is more often the domain of the social sciences. Some of the participants felt that these materials might not necessarily meet the needs of researchers in the humanities. With further discussion, the groups concluded that as both disciplinary fields adopt both quantitative and qualitative methodologies, it made more sense to think



of potential audiences as being either quantitative or qualitative researchers and consider how these materials will meet the needs of both.

Recommendations:

1. Consider which groups will benefit from these materials and make this clearer in the materials.
2. Adding some examples of outputs based on qualitative data

I designed the materials originally with quantitative researchers in mind, and the participants agreed that the training would address the needs of quantitative researchers, as the content specifically addresses issues surrounding the sharing of numeric data sources. Thus, any researcher carrying quantitative analyses, regardless of their academic discipline, could benefit from these materials.

For qualitative researchers, the picture is less clear. Some analysis software packages such as NVivo produce quantitative data, in which case these materials would retain some value. For those employing other qualitative analysis methodologies which do not produce quantitative data, the early modules would still be relevant, but the modules on disclosure risk will not be directly relevant. There is no immediate solution here. Currently research is still ongoing on how disclosure control can be applied to qualitative research output, and a key outcome from this project is that the materials are to be further developed once this research bears fruit to include disclosure control examples for qualitative research outputs.

## 2. The inclusion of legalisation information

The second theme focused on the role of legislation in data access which is covered in module 4. This divided opinions and prompted lively discussion among the participants with experience in the development and delivery of this kind of training. There were two schools of thought: one supporting the inclusion of information about legislation as a means of highlighting the importance of data protection and the role that TRE procedures have in ensuring that researchers do not breach their legal responsibilities. The other perspective was that providing information about legislation is unnecessary as researchers should be encouraged to feel a sense of community and responsibility towards working with TRE staff. This discussion mirrors previous discussions that I participated in as part of the SRT Expert Group, set up in the UK to discuss and steer the Safe Researcher Training courses.

No firm recommendation emerged in this area within the small group discussions. Within the scope of developing canonical materials that are adapted for a wide range of TREs, I recognise that both perspectives are equally valid with the decision on how much information on legislation to include determined, at least in part, by the target audience. As a project, we decided to keep this content in the canonical materials to give the individual TREs the option to decide how much legal information to include and whether to include this content in a presentation or in a supplementary handout. For those wishing to exclude the content on this topic, they could simply remove the complete module, without the need to carry out further edits to the rest of the materials.

## 3. The length of training courses based on these materials



The third theme was around how long courses based on these materials would take to deliver. Among the participants in group 2 were several experienced 'Safe Researcher Training' trainers and they along with myself were able to give guidance on estimated course lengths of between 2.5 and 5 hours depending on the mode of delivery. The conclusion of group 2 was that the length of the course is also determined in a large part by the level of prior experience of the participants, how communicative they are, and the delivery style of the trainers.

Recommendation:

The key recommendation from the discussion in group 2 was that guidance on the length of the course should be included in the training materials. This will be added in future iterations of the materials.

#### 4. Delivery modality of the training materials

Another key theme to emerge in the discussion groups was how to deliver the training, i.e., which delivery modality would be the most effective method. I designed the materials to function as a taught course, that offered the possibility of developing either an in-person or virtual training. There are benefits to both delivery modes. Delivering training in person allows trainers to build up a good relationship with the researchers who they will be supporting. This can be particularly useful when it is necessary to discuss potential problems with researchers. However, in-person training is time and cost intensive. For example, at the UK Data Service, the SecureLab team delivered Safe Researcher Training with two trainers approximately every three weeks in London. The burden of traveling was also felt by researchers, who may have had to travel some distance to attend the course.

Virtual training offers a cheaper, less resource intensive option, but in group 1, participants felt that even virtual training might be too resource intensive for smaller TREs. They felt that the modular design would allow the content to be adapted fairly easily to an online self-study format, should that be the preferred delivery mode. This may be an attractive option for services that do not have the resources to run regular taught training events, either in-person or virtually.

Recommendation:

Consider adding further adaptations or guidance to the materials to aid TREs who wish to develop their training as an E-learning program.

As part of the SSHOC project we were not able to fully develop further iterations of the slides themselves, so this recommendation remains outstanding. However, we believe that the core content could be used and built upon to develop a self-study course material. For example, more information would need to be added to the slides, but this can easily be taken from the information I included in the delivery notes. Depending on resources, TREs could consider also producing short video recordings for some of the content to try to engage more active engagement with the content.

#### 5. Delivery notes for trainers

The final theme to emerge was around how trainers could be supported in deliver training, especially those new to Safe Researcher training. Some participants felt that although I included many delivery notes in the PowerPoint slides, trainers new to the topic area would still need further guidance. They

also felt that some guidance for TREs on deciding what content to include and what delivery mode to opt for would be helpful. To address these issues, the participants felt that a separate trainers guide would be a good addition to the materials.

**Recommendation:**

A separate supplementary guide for trainers should be developed and included in the training materials. This guide should include not only delivery notes for trainers, but also some guidance on some of the issues discussed in this section.

Following the workshop, I produced a supplementary guide for trainers which starts with an overview of the purpose of the training materials, and then guidance on to TREs on understanding their audience and how to decide on the delivery mode for their training courses. The remainder of the guide provides slide-by-slide guidance on the overall purpose of the slide and some ideas for key points to discuss. These notes are not intended to be a verbatim script but should guide trainers to the content and its purpose. This guide has now been added to the set of materials and made available online along with the Powerpoint slides. As it stands, we have not had chance to gain feedback from TREs and trainers, however, we hope to be able to assess the guides effectiveness in addressing these issues in the future, once we have some use cases for the materials.

### **Concluding comments and next steps**

The development of a new set of canonical Safe Researcher training materials is an important step in expanding sensitive data access. The materials were well received by those attending the workshop underlined the development of the materials as an important development in supporting TREs to implement training for researchers applying to access sensitive data. The consensus in the small group discussions was that the comprehensive breadth of information included in the materials means that TREs would be able to adapt them to their needs with relative ease. The materials are now freely accessible via the SSHOC project Zenodo account, and to date have been downloaded 128 times (Wiltshire, 2021a).

Overall, it was widely agreed that these canonical training materials would be of great benefit to the international TRE and research communities and could play a role in driving consistent standards of training across different organisations and countries. This is particularly important with the move towards opening up international data sharing where consistent data governance structures are vital. But this work has also highlighted the need to conduct further research into Safe Researcher training so that we can move beyond relying on anecdotal evidence.

Since the completion of the SSHOC project, there have been two exciting developments. First, several TREs in Germany have now expressed interest in developing Safe Researcher training for their own services. As a result, I am utilizing and adapting the SSHOC canonical training materials as part of a new collaborative project in Germany, the 'Accrediting Safe Use of Research Environment (ASSURED) project. We are currently working to develop a modular E-learning training program for TREs across Germany, which initially will offer training pathways

for researchers wishing to access sensitive data via TREs and for staff working in TREs wishing to further develop their knowledge and improve their career development opportunities<sup>9</sup>. This will be the first use case of these materials and will offer over the coming months a chance to more formally assess their efficacy.

The second development was a consensus among workshop participants that the opportunity to come together to discuss key aspects of secure data access and use, and to benefit from each other's knowledge, was a valuable experience for those working within TREs. This spirit of cooperation and engagement formed the basis for the new International Secure Data Facility Professionals Network (ISDaNet) (Lichtwardt, Wiltshire and Bishop, 2022). Since the SSHOC workshop, this new Network has met biannually, bringing together TRE professionals from across the globe to discuss different aspects of secure access to sensitive data.

With this spirit of cooperation and collaboration, it is hoped that the endeavours to build consistent and comparative practices across the global TRE community will further advance the move towards wider, more equitable access to sensitive data.

## References

Bishop, L., Broeder, D., van den Heuvel, H., Kleiner, B., Lichtwardt, B., Wiltshire, D. and Voronin, Y. (2022). 'D5.10 White Paper on Remote Access to Sensitive Data in the Social Sciences and Humanities: 2021 and beyond (1.0)', *Zenodo*. Available at: <https://doi.org/10.5281/zenodo.6719121>.

Desai T. and Ritchie F. (2010) 'Effective researcher management', *Work session on statistical data confidentiality 2009*, pp. 1-11. Available at: <https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2009/wp.15.e.pdf>.

Desai, T., Ritchie, F. and Welpton R. (2016) 'Five Safes: Designing data access for research'. *Economics Working Paper Series 1601*, pp. 1-27. Available at: <https://www2.uwe.ac.uk/faculties/BBS/Documents/1601.pdf>.

Lambert, D. (1993) 'Measures of Disclosure and Harm', *Journal of Official Statistics*, 9(2) 313-331. Available at: [https://www.researchgate.net/profile/Diane-Lambert/publication/2337915\\_Measures\\_of\\_Disclosure\\_Risks\\_and\\_Harm/links/58486b7a08ae95e1d1665c22/Measures-of-Disclosure-Risks-and-Harm.pdf](https://www.researchgate.net/profile/Diane-Lambert/publication/2337915_Measures_of_Disclosure_Risks_and_Harm/links/58486b7a08ae95e1d1665c22/Measures-of-Disclosure-Risks-and-Harm.pdf).

Lichtwardt, B., Wiltshire, D. & Bishop, E.L (2022). 'D5.12 International Secure Data Facility Professionals Network (ISDFPN)', *Zenodo*. Available at: <https://doi.org/10.5281/zenodo.6583379>.

Social Science and Humanities Open Cloud (2021) 'Workshop Blog; Workshop notes: Developing canonical training materials for secure data access facility professionals', *SSHOC*, 30 September. Available at: <https://sshopencloud.eu/news/sshoc-workshop->

11/12 Wiltshire, Deborah (2024) Developing canonical 'safe researcher' training materials for trusted research environments, *IASSIST Quarterly* 48(1), pp 1-12. DOI: <https://doi.org/10.29173/iq1093>

[notes-providing-canonical-training-materials-secure-data-facility-professionals](#).

Wiltshire, D. (2021a). 'D5.20 Training materials of workshop for secure data facility professionals (v1.0)', *Zenodo*. Available at: <https://doi.org/10.5281/zenodo.5638596>

Wiltshire, D. (2021b). 'SSHOC Workshop: Providing canonical training materials for secure data facility professionals', *Zenodo*. Available at: <https://doi.org/10.5281/zenodo.5541587>.

Wiltshire, D. and Alvanides, S. (2022). 'Ensuring the ethical use of big data: lessons from secure data access', *Heliyon*, 8 (2), pp. 1-6. Available at: <https://doi.org/10.1016/j.heliyon.2022.e08981>

Woollard, M., Lichtwardt, B., Bishop, E.L. and Müller, D. (2021). D5.9 Framework and contract for international data use agreements on remote access to confidential data (v1.0)'. *Zenodo*. Available at: <https://doi.org/10.5281/zenodo.4534286>

---

## Endnotes

<sup>1</sup> Deborah Wiltshire, GESIS-Leibniz Institute for the Social Sciences, Unter Sachsenhausen 6, 50667 Cologne, Germany, [deborah.wiltshire@gesis.org](mailto:deborah.wiltshire@gesis.org); Orcid 0000-0001-6533-2426

<sup>2</sup> Examples of a Safe Haven or Safe Room include the UK Data Service SecureLab (<https://ukdataservice.ac.uk/help/secure-lab/what-is-securelab/>) and the Secure Data Center at GESIS (<https://www.gesis.org/en/services/processing-and-analyzing-data/analysis-of-sensitive-data/secure-data-center-sdc>) [Accessed 01/08/2023].

<sup>3</sup> Become an accredited researcher, 2023 <https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/secureresearchservice/becomeanaccreditedresearcher#full-accredited-researcher-under-the-digital-economy-act-2017-dea> [Accessed 30/07/2023].

<sup>4</sup> <https://sshopencloud.eu/about-sshoc> [Accessed 08/08/2023].

<sup>5</sup> <https://eosc-portal.eu/> [Accessed 29/07/2023].

<sup>6</sup> Wiltshire, D. 2021. D5.20 Training materials of workshop for secure data facility professionals (v1.0). *Zenodo*. <https://doi.org/10.5281/zenodo.5638596>

<sup>7</sup> JamBoard is a free interactive Whiteboard from Google which allows participants in virtual classrooms to collect ideas and work collaboratively: <https://jamboard.google.com/> [Accessed 11/08/2023].

<sup>8</sup> See SSHOC workshop: Developing canonical training materials for secure data access facility professionals; <https://www.sshopencloud.eu/events/sshoc-workshop-providing-canonical-training-materials-secure-data-facility-professionals> [Accessed 11/08/2023].

<sup>9</sup> This project is a collaborative endeavour by the Secure Data Center at GESIS, the German Human Genome-Phenome Archive and Berd @ NFDI. It does not yet have a dedicated website but a presentation outlining the project (using its original project name SaRAS) can be found online: Wiltshire, Deborah. (2023, March 9). Safe Researcher Accreditation System. *Zenodo*. <https://doi.org/10.5281/zenodo.8238606>