

VML USER TRANSITION

A public document for VML researchers wishing to transfer to the Secure Data Service

Scope

The aim of this document is to provide existing users of the ONS Virtual Microdata Laboratory (VML) with information about how they can transfer to the Secure Data Service.

It is not within the scope of this document to explain how the Secure Data Service works. For questions about the Secure Data Service researchers are advised to contact the Service directly. (Details are at the end of this document).

Introduction

The establishment of the Secure Data Service represents a significant change to how confidential business, economic and social data can be accessed by the research community in the UK.

Funded by the Economic and Social Research Council, and supported by the Office for National Statistics, the Secure Data Service will provide more convenient access to researchers who wish to analyse these data.

The ONS Virtual Microdata Laboratory (VML) has been a valuable research service that has attracted a huge demand, particularly for the business data collections it holds. However, the need to travel to government offices is a significant barrier to use of the data.

The Secure Data Service aims to remove that barrier for academic users, while still providing a secure and flexible research tool. Members will be able to log in to the Secure Data Service from their institution. The service is available 24 hours a day, 365 days a year, through a familiar Windows™ desktop environment. Users will have access to the data they have applied to use, and familiar statistical software such as Stata and SPSS.

The ESRC currently subsidises academic access to the VML, by covering the access charges. The daily charge to use the VML is currently £52 + VAT. This will be reviewed after the Secure Data Service transition is complete (when ESRC subsidies for academic research in the VML is discontinued), but is likely to be considerably higher in the future. Academic users who continue to use the VML when the Secure Data Service could be used instead will now be responsible for paying the daily charge.

The Secure Data Service is a more convenient environment for researchers, and will be free of charge to academic researchers. Although many researchers in Higher Education institutions (universities) and Research Institutes across the UK have used the VML – and will have projects they are currently working on – it is expected that the combination of convenience and cost will encourage researchers to transfer work to the Secure Data Service wherever possible.

This document is aimed at those researchers who currently use the VML and wish to transfer to the Secure Data Service. By reading this document, we hope that researchers will understand:

- How information on the transfer will be communicated
- The ‘Seven Steps to Transferring’
- What files they can and cannot transfer to the Secure Data Service
- How their VML files will be stored
- How Postcodes and Grid-References will change
- Why they might wish to use the VML in the future
- Where they can find the help and support needed to successfully transfer

Communication with Users

Information about the Secure Data Service will be available by visiting the website securedata.data-archive.ac.uk. VML users will also be given information about the transferring to the new Service via the VML Quarterly Bulletin. In addition, researchers can contact the Secure Data Service directly at securedata@data-archive.ac.uk or 01206 87 4968.

To use the Secure Data Service, all VML researchers will be required to attend and pass through the Secure Data Service training seminar (a half-day event at the researcher’s institution or another venue nearby). A number of courses have already been set up, all current and future courses will be advertised via the VML and the SDS website .

During the seminar, researchers will be informed in more detail about the procedures for transferring their work to the Secure Data Service, based on guidance from the VML.

Seven Steps to Transferring

We have identified seven steps that researchers pass through in order to successfully transfer their project from the VML to the Secure Data Service. These are:

1. **Attendance at a Secure Data Service training seminar** (this is required for anybody who wishes to use the Secure Data Service, even if they have already used the VML)
2. After successfully completing the Secure Data Service training seminar, **the Secure Data Service and the VML will exchange the paperwork relating to the researcher(s) project(s)** (for example the Approved Researcher forms originally submitted to the VML). A researcher who wishes to transfer live projects will not have to apply to become an Approved Researcher again (unless they begin a new project).

3. **The researcher should then book to use the VML**, and arrange the files they wish the VML to transfer to the Secure Data Service (for more information about the types of files they can transfer, please see [‘What files can be transferred?’](#) and [‘Making the Move’](#))
4. When the Secure Data Service has received information about a researcher’s VML project, an account will be created for them. The Service will transfer the users’ files to that account, after the transfer of files has been approved by the VML.
5. **Each researcher must then register their project with the Secure Data Service** through the Economic and Social Data Service website (a link to the website containing instructions will be provided). Researchers then order the datasets they require.
6. **When the Secure Data Service has received the order from the researcher, the data are deposited into their account.**
7. **The researcher should return their signed User Agreement to the Secure Data Service** (this relates specifically to each project and should be signed by the Head of Department). The researcher will then receive their Secure Data Service member login details.

What files can be transferred?

The VML has agreed that all researchers may transfer syntax files that they have written to create data files and execute commands for their analysis. In general, only syntax files will be transferred: users’ data files will **not** be transferred.

Users may apply to transfer ‘user-created datasets’ and these requests will be dealt with on a case-by-case basis. These are not ‘working files’ that researchers should easily be able to reproduce with good syntax. Instead, these are datasets that have been developed and established using sources of ONS data, and which are used routinely by researchers on multiple projects.

Researchers will also be allowed to transfer other ‘miscellaneous’ documents, such as Word, Excel, PowerPoint slides, PDFs etc. These may be documents that contain preliminary presentations, commentaries on datasets or results, etc. They should not contain data, either extracted from an ONS data file, or from external data.

The transfer of these other documents will be at the discretion of the VML team, who will require evidence that the documents are of significant value and difficult to reproduce.

Overall, users are advised to consider the Secure Data Service as a ‘clean sheet’, rather than a duplicate of their existing environment.

During transition, researchers needing time in the VML to clean up their project areas (as opposed to carrying out research) will not face the daily VML access charge.

Researchers should note that any files that they wish to have released to them outside the secure confines of the VML or the Secure Data Service will be subject to normal Statistical Disclosure Control checks, and that both the VML and the Secure Data Service reserve the right to refuse such requests subject to their discretion.

Making the Move

Each researcher will need to make a booking for the VML. During this visit, the researchers should organise the files they wish to request that the VML transfer to the SDS on their behalf.

These files should then be placed in the output drive, under the user name, in a subfolder called 'transfer files – project code xxxxxx'. To request the transfer, a template has been created, in the folder 'clearance forms'. This will be picked up by the VML team in the normal way.

For requests to transfer data files (see previous section) a separate template has been created. Data files should not be copied to the output drive.

Storage of files in the VML

After a researcher has successfully transferred to the Secure Data Service, the user's VML account will be closed. However, the user's files will be retained for at least two years as a backup in case of a later need to identify more files.

New geography

The detailed geography currently available in the VML data poses a risk of identification (for example, from postcodes). As part of the move to the SDS, the geographical information in both the VML and SDS will change.

The VML and SDS data files will no longer contain real postcodes. Instead, they will be replaced by a pseudo-anonymised postcode which maintains the hierarchical structure of the postcode but without any real values.

All files containing postcodes (including lookup tables) will be recoded using the same anonymisation process, so that all references to postcode are consistent. At the same time, geo-markers (eastings and northings) will be edited to prevent direct identification; again, all files in the VML and SDS using the information (for example, shape files for county boundaries) will be edited for consistency.

The impact of these changes should be minimal. Postcodes and geo-markers will retain the statistical characteristics of the current variables; and within the VML and SDS they will be consistent. In fact, it will be easier to exploit the hierarchical structure of the postcodes as the pseudo-anonymised variables have a fixed length for each part of the postcode. However, they will not be consistent with any external detailed geography.

Be aware that any researcher found attempting to reconcile internal and external geographies will be in breach of procedures and subject to disciplinary action.

Researchers wanting to, for example, bring in additional shape files will therefore need to ensure that these are appropriately recoded.

More information about these changes can be found in the document "Using the VML-Secure Data Service Pseudonymised Postcodes" and researchers should contact the VML team at maus@ons.gsi.gov.uk if there are any further questions

Future role of the VML

Depending upon the data researchers wish to use, there may be a reason to continue using the VML.

For example, if researchers require access to highly sensitive data that can only be analysed in a Government environment, then the VML may be able to provide assistance. They can work with the Administrative Data Liaison Service (www.adls.ac.uk) to provide data that can only be accessed in a Government setting.

Researchers who access the CAMS (Controlled Access Microdata Sample) data, or the Longitudinal Study, will need to continue to access those data through the VML.

Researchers may also wish to analyse their own data together with data in the Secure Data Service. For example, if a researcher has a survey of companies that they wish to analyse in the Secure Data Service, the VML can arrange for the sample to be matched to the anonymous IDBR references, which can then be used in the Secure Data Service.

Furthermore, the VML plans to expand by providing access to civil servants in other Government Departments. A researcher working with a civil servant can be supported by the two services.

Support and contact information

Further information about transferring from the VML to the Secure Data Service will shortly be available at a new Secure Data Service website, which will have a list of frequently asked questions (FAQs) to help researchers.

In the meantime, please contact the Secure Data Service team directly:

Tel: 01206 87 4968

Email: securedata@data-archive.ac.uk