“Remote Access to European Microdata”

Maurice Brandt (maurice.brandt@destatis.de)

Keywords: Remote Access, Confidential Microdata, Improvement of Access, Science

1. INTRODUCTION

To improve the access to European microdata for scientific purposes the ESSnet-project “Decentralised and Remote Access to Confidential Data in the ESS” (DARA) was conducted. The aim of the ESSnet-project DARA was to establish a secure channel from a safe centre within a National Statistical Institute (NSI) to the safe server at Eurostat, so that researchers can use confidential EU microdata in their own Member States without travelling to Luxembourg.

The ESSnet DARA-project team has defined a concept of technical implementation and safety requirements for a European remote access system. The concrete task of participating NSIs was to provide a secure channel to guarantee access for data users to the central node and also to provide service and IT-support for the researchers on the local national level.

The project team has drafted a handbook with descriptions and guidelines for NSI staff and researchers and an accreditation system for access facilities. For a proof of the concept and feasibility, the project team has implemented a remote access pilot with 6 access points in 5 countries in Europe.

2. METHODS AND WORK COMPLETION

One main task was the preparation of the implementation of a pilot infrastructure from a technical point of view. At the beginning of the ESSnet DARA project, the team developed security requirements and user demands for a European Remote Access System.

One crucial topic that was investigated was the detail of the connection to the central system. The safe connection between the European Commission (EC) and the NSIs in the Member States is guaranteed by a safe network named “Secure Trans European Services for Telematics between Administrations” (sTesta) based on a private network. Only institutions that are part of the private network are able to connect to the CITRIX server within the network of the EC. This means that the NSIs have to join the sTesta network before they are able to establish a connection. There is also a login and password required to access the working platform. On the one hand it is a higher burden for the NSIs. On the other hand the whole system itself is more secure because it is not possible to connect to the CITRIX server from any other location outside the network. During the preparation phase of the pilot, the method for connecting to the Citrix server at the EC was investigated. Until midterm of the project there were two NSIs connected via sTesta network. Other NSIs had difficulties to connect via sTesta because this part of the IT infrastructure was outsourced to an external company. This also made the connection method much more expensive as the external companies were charging monthly fees for maintaining the connection. For various reasons it is not widely used by the NSIs in the
MS. The method for a secure connection was not very promising during the first half of the project. This is why further investigation on a secure channel via VPN has been conducted. The aim was to find a secure mode of connection between MS and Eurostat that can be used by all MS.

To test and evaluate the usability of the pilot, a testing plan was drafted. The aim of the proof of concept was to set up the testing phase of the pilot of a trans-border Remote Access for each country involved in the project.

The pilot was designed for a European system, but it was built from an existing solution in France. The project has set up a new infrastructure based on the requirements defined at the beginning, like user needs, security requirements, roles, workflow specifications, etc.. The conclusions are that the proof of concept demonstrates that all requirements defined are relevant and a solution that fulfils all these requirements can be implemented considering the security of the system, the usability for the users, and the decentralised management in each country.

The pilot shows that a European Remote Access System can be implemented in safe conditions. To establish a connection, only a broadband internet connection is needed. First the IP settings and the Proxy-server need to be configured in the so called DARA Box. A registered person like a Support Officer or researcher can login with a smartcard and fingerprint reader. Then they can access the central server via an encrypted and secure connection to a virtual machine using a familiar windows desktop that has statistical editing and analysis software with MS Office programs.

3. RESULTS

After the experience of the ESSnet DARA, the project team can give a clear recommendation on a European microdata access infrastructure. After first tests the alternative pilot looks very promising in terms of security, costs, management, deployment and delegation in a European context and user friendliness. Therefore the recommendation is the implementation of the DARA pilot with Eurostat as the central node. It can be also a system that fulfills all defined user and security requirements with Eurostat as centre, where the DARA pilot is an example that has been tested successfully according to the specifications. The first accreditations of Access Facilities, starting with NSIs, will show how the realisation of the Regulation (EU) No 557/2013 works in practice. For the near future it is advisable to build up a circle of trust and competences for microdata access for scientific purposes in Europe.

There are also lessons learned that lead to recommendations which cannot be given for a production phase of a Remote Access System. Such a system has very specific requirements and should be implemented in a dedicated environment and not on a shared platform with other services. The sTesta connection is appealing, but for practical reasons it should be avoided for a real implementation because it is not widely used and an extension to other access points for a long term vision is doubtful.

During the project there have been requests from researchers in Europe, especially from the OECD, Portugal, Hungary and the UK for detailed microdata on the Labour Force Survey and other datasets. The interest was huge to use those data from location near their own research institutions. This shows that there is demand and utilisation for this system and it should be also expanded to other EU statistics and countries.
The complete final project report is accessible on cros portal under the following link:

And a summary paper of the ESSnet DARA project is available under:

4. CONCLUSIONS

In conclusion, the DARA pilot shows, as a proof of concept, that it is possible to set up a European decentralised Remote Access to confidential microdata in excellent conditions considering:

- A high level of security with a strong authentication method and a leak-proof infrastructure prohibiting data files extraction.
- A good usability for researchers.
- A flexible system management that allows delegation to local Safe Centre for enrolment, project creation, output checking, etc.
- An easy method for the deployment and the installation.
- A cost-effective solution for both central node and remote sites (no strong local security IT requirements).

All this was possible due to the specifications produced, like user requirement, handbooks, workflows, IT security requirements and user needs. The proof of concept validates all the concepts defined in this way, and furthermore helps to produce the cost analysis study.

The extended study, produced by the project, would be very useful for the design of the real implementation as well as for the production phase (handbook, workflows, requirements, etc.). The benefits for this European microdata access system are that secure data server and devices for thin clients in the Safe Centres could be provided by the central node so that there are no investments for IT equipment necessary for the MS planning to join this system. No microdata will be transferred to another MS, only the access will be granted from another accredited Access Facility whereas the microdata itself will remain at the secure servers inside Eurostat. This system needs to be affordable to maintain it over the years and to build a sustainable solution which can be also used in the future.

Only if the mode of microdata access is secure and user friendly the data can be used by European researchers for their analysis. This will contribute to a better understanding of processes and developments in Europe and can help to find best practise examples that can improve conditions in all Member States. In a context of a “European society” only an evidence based assessment of the situation with empirical microdata can lead to realistic measures for Europe. Furthermore, actions which have been implemented in Europe can be evaluated and adjusted if high quality EU microdata is available for researchers who are working and experienced in the field of empirical European studies. All this can contribute to the improvement of social and living conditions in Europe.
REFERENCES


