

Improving access to social science data

★ Archiving data for social science research and creating an infrastructure for accessing these resources for a wide range of researchers in various locations is a task that the CESSDA Preparatory Phase Project (CESSDA-PPP) is charged with, as **Kevin Schürer** explains

The Council of European Social Science Data Archives (CESSDA) is an umbrella organisation for social science data archives across Europe. Since the 1970s its members have worked to improve access to data for European researchers and students. CESSDA currently provides access to 25,000 data collections, delivering more than 70,000 data-sets per annum and acquiring a further 1,000 data collections each year.

CESSDA research and development projects and 'Expert Seminars' enhance the exchange of data and technologies among data organisations. Data collections include sociological surveys, election studies, longitudinal studies, opinion polls, and census data. Among the materials are cross-national comparative data such as the European Social Survey, the Eurobarometers, and the International Social Survey Programme.

The CESSDA Preparatory Phase Project (CESSDA PPP) is a 2.7 million-Euro project aimed at developing the CESSDA research infrastructure (RI), and is a direct result of the CESSDA RI being identified by the European Strategy Forum on Research

Infrastructures Roadmap (ESFRI) exercise as a research network of excellence. It focuses on tackling and resolving a number of strategic, financial and legal issues in order to ensure that European social science researchers have access to, and support for, the data resources they require to conduct research of the highest quality, irrespective of the location of either researcher or data within the European Research Area.

Creating structure

UK-based CESSDA President and project PI, Prof K Schürer, says this project is mainly about creating a new formal legal entity out of what has up until now been an informal grouping, involving organisations in different European countries. "Each of these," he said, "has common goals of preserving and sharing data from social sciences, and making them as readily available and as useful for users as possible – negotiating access with data producers to increase the quantity and quality fund of data available to social scientists.

"The situation at the moment is that the take-up, or embracing, of data archiving, is different in different countries. A significant

part of this project is about promoting the role of the data archive in countries where they exist, but don't have the total buy-in from people at the top who can provide the funding needed to bring some of the less-well resourced organisations up to speed with developments that are happening elsewhere. This will give them the opportunity to capitalise on what others are doing - thereby effecting knowledge and skills transfer between the older and newer, the larger and smaller, organisations.

"Our goal has always been to share expertise within the community as a whole and to raise professional standards and skills. Yet imbalances between countries in terms of the resources available is still a problem, and in some cases a barrier to the pace of development. Thus in planning the new organisation, in order to create fully interoperable systems, it will also be important to create a more equal, accountable and sustainable funding base."

The idea of CESSDA organisations collecting data and 'archiving' them has developed in recent years, he said. "We pay more attention to data preservation now rather than just thinking of it as

archiving. Archiving is often thought of as a very static thing – you put something into a cupboard and then forget about it, and if somebody does want it in 100 years time then OK, they could leaf through it and glean some information. However, this can't work for digital materials.

"Instead, we actively curate, making sure the data are held in a format that is going to minimise possible loss. For example, so they're not dependent on software or hardware, so that in 10 years' time when that right software or equipment or storage media no longer exist or work, you haven't lost the data.

"At the same time what is also important is to work in partnership with researchers

have facilities for making tabular material available, or they can direct researchers or journalists to an organisation that can. It may be the organisation that produced the data-set in the first place or it may not, or it could be done internally. But there are tools on the web that will allow that to be done easily and regularly, and I think it will happen more in the future than perhaps it does now. There is new legislation across Europe and a growing movement to make publicly funded data available. This will ultimately reduce the restrictions in making these materials available. Obviously intellectual property will remain important, and certainly data

“ **We pay more attention to data preservation now rather than just thinking of it as archiving. Archiving is often thought of as a very static thing – you put something into a cupboard and then forget about it, and if somebody does want it in 100 years time then OK, they could leaf through it and glean some information** ”

and say 'you have great resources here and here are some ways their value can be maximised and sustained over time.' This is a kind of training process for researchers. We do a lot in working with and encouraging researchers to think about the ethics of data use and data collection and about data sharing. For these reasons, in the UK, for example, the Economic and Social Research Council requires any researcher who is funded by them to talk to the UK Data Archive about making the data they collect available once the funding award ends. Similar requirements exist in other countries, yet more still needs to be done.”

Wider data access

One of the results of the CESSDA-PPP project will be to make the data infrastructure available to the wider audience, as long as those using them are prepared to accept any data-protection conditions associated with usage. "Increasingly people on the street – journalists in particular – are interested in getting information from a data-set," Hilary Beedham, the Project Manager for the CESSDA – PPP said. "So, some archives

protection matters and data subject rights must always be paramount.”

The future of the CESSDA-PPP project will see the organisation reach out from its European boundaries to other countries. Hilary Beedham would like to see all European countries participating – “if we don't, then the infrastructure won't be as effective” – but she hopes the projects aims and ambitions will be accepted on a global scale. “All of the plans that we have for the new organisation are aimed at allowing it to be as inclusive as possible,” she said, “so we're talking with the Australians, and the Americans are also very interested in this work so we will be extending this beyond Europe. Certainly there will be more data, because each new member will bring new resources. We assume that the members will be organisations that have an interest in making data available.

“We will be talking to organisations like the World Bank to tell them about the potential for collaboration in making their data more easily available for researchers. There's infinite potential for further expansion. It will be a long time before the organisation becomes saturated.” ★

At a glance

Full Project Title

Council of European Social Science Data Archives Research Research Infrastructure (CESSDA)

Project Objectives

Our goals are to: Transform CESSDA into a legal entity; To extend the CESSDA network; Promote member organisations as places of deposit for publicly funded data collections thus extending their collections for the benefit of the ERA; Provide a one-stop shop for data discovery, access, analysis and delivery; Increase the quality of available data; Create a more dynamic knowledge management oriented web, where knowledge-products create bridges between text in scientific journals and the underlying data; Establish minimum requirements for CESSDA membership and agreements between CESSDA partners and data-related organisations outside the defined Research Infrastructure.

Information on CESSDA-PPP's sister project Dariah, appearing in edition 17 of *Projects*, is available at: www.dariah.eu/

Contact Details

Project Coordinator,
Professor Kevin Schürer
Director, UKDA/ESDS & President,
CESSDA
UK Data Archive, Wivenhoe Park,
University of Essex, Colchester, CO4
3SQ, UK
T: +44 (0)1206 872570
E: beedh@essex.ac.uk
W: <http://www.data-archive.ac.uk/>

Kevin Schürer



Project Coordinator

Kevin Schürer is the UK representative for the European Strategy Forum for Research Infrastructure (ESFRI) working group in Social Science and Humanities and current President of the Council of European Social Science Data Archives. He co-ordinates the CESSDA PPP, has been Director of the UK Data Archive since 2000 and holds a Professorship in History at the University of Essex.