Synergies for SERISS

Maintaining state-of-the-art research infrastructures is increasingly complex and costly, requiring the integration of different tools and data sources, as well as extensive transnational collaboration. The European Commission's ongoing support for infrastructure development through cluster projects for a variety of scientific disciplines recognises this fact. Crucially, these cluster projects facilitate collaboration and infrastructure building not only within the physical or biological sciences but also within the social sciences. Just as investment in physical infrastructure such as telescopes is necessary to better observe the physical world, so investment in social surveys and infrastructures is necessary in order to better observe society and address key societal challenges such as migration, climate change and rising inequality.

Synergies for Research Infrastructures in the Social Sciences (SERISS), funded by the European Commission as part of its Horizon 2020 programme – under Grant Agreement No 654221 – is a four-year collaboration (2015-2019) between the three leading European research infrastructures in the social sciences – the European Social Survey (ESS ERIC), the Survey for Health Aging and Retirement in Europe (SHARE ERIC) and the Consortium of European Social Science Data Archives (CESSDA AS) – and organisations representing the Generations and Gender Programme (GGP), European Values Study (EVS) and WageIndicator. The project exploits synergies, fosters collaboration and is developing shared standards to equip national and

The value of social surveys for drawing inferences about public behaviour and attitudes relies on generating data which is both representative of the population under investigation and equivalent across countries
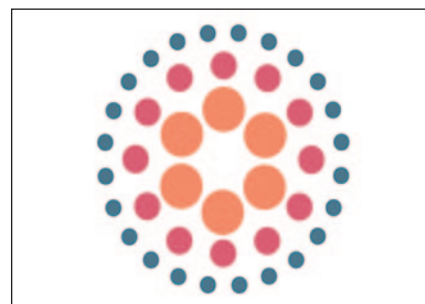
European policy makers with a solid base of the highest-quality evidence on people's attitudes, experiences and behaviour.

This article outlines how SERISS aims to strengthen Europe's social science evidence base by addressing key challenges facing cross-national data collection, breaking down barriers between research infrastructures and embracing the future of social science via new forms of data collection.

Cross-national research

The value of social surveys for drawing inferences about public behaviour and attitudes relies on generating data which is both representative of the population under investigation and equivalent across countries. This is not an easy task. Achieving a representative sample is complicated by restricted access to sampling frames, the varying nature and quality of sampling frames, and a climate of lower response rates. SERISS will address these challenges in a number of ways. First, by mapping sampling practice across European surveys with the aim of exploiting synergies and economies of scale and possibly pooling sampling resources in future. Second, by assessing potential for exploiting pre-existing administrative data to better understand and overcome survey non-response. We will also examine the feasibility of including institutional populations in general population sample surveys.

Achieving equivalence in cross-national surveys relies, amongst other things, on high quality translation procedures. There are three key areas of work under SERISS to advance translation best practice. The first involves comparative testing of different questionnaire translation approaches to provide empirical evidence on best practices for country-specific adaptations in survey translation. The second explores the feasibility of applying computational linguistic methods to survey translation. The third involves the updating of an online Translation Management Tool (TMT) already used by the SHARE survey for use by other surveys.

Breaking down barriers

Even the most sophisticated surveys still tend to rely on basic tools and processes, resulting in wasted or duplicated effort as well as scope for important information to be lost and inconsistencies to occur. SERISS is developing a number of cutting-edge tools to streamline and harmonise the implementation of the survey lifecycle and to facilitate more efficient, standardised and transparent working both within and across cross-national survey infrastructures.

Tools being developed include an online Question Design and Documentation Tool (QDDT) and Question Variable Data Base

The digital revolution opens up opportunities for the social sciences in the shape of new forms of data and data collection methodologies

(QVDB) which, together with a TMT, will generate and store metadata about different stages of the survey lifecycle. Other tools being developed include a web-based survey management portal to provide a virtual collaborative workspace for multiple stakeholders on international projects and a data harmonisation platform, which will facilitate more efficient and well-documented output harmonisation by providing a platform via which data users can deposit and receive feedback on their harmonisation routines.

Many of the tools will be DDI-compatible in order to facilitate the efficient generation and sharing of survey metadata across different survey infrastructures. The tools mean that reliable and comprehensive metadata will be available not only to the primary data producers but also to secondary users interested in data analysis or replication on other surveys.

Embracing the future

The digital revolution opens up opportunities for the social sciences in the shape of new forms of data and data collection methodologies. It also raises new challenges including issues around access, quality and data security. One important strand of work under SERISS centres on exploring the legal and ethical challenges associated with harnessing new forms of data. Mapping exercises are currently underway to explore ways in which two increasingly important sources of data – social media data and administrative data – are being used by social scientists and to identify potential barriers to their access and use. Findings from the mapping exercises will be considered alongside the new European Data Protection Regulation to develop best practice guidelines to promote the reuse of these new forms of data by social scientists.

Data collection is increasingly moving online. However, cross-national research using web surveys remains largely dependent on data generated by convenience samples and/or ex-post harmonisation of country-specific studies. A lack of representativeness and/or differences in methodology across countries can damage the scope for meaningful cross-national comparisons using such data. SERISS will test a proof of concept for a co-ordinated cross-national probability-based web panel. A small web panel will be established in three countries (the UK, Slovenia and Estonia) by drawing respondents from among those participating in the main face-to-face wave of the European Social Survey in 2016. They will then be followed up every other month for one year.

The web panel, CRONOS, provides an opportunity to explore different aspects of panel design and administration, including strategies for respondent recruitment and retention, and measurement issues associated with online data collection, including strategies to overcome 'bad' respondent behaviour such as straightlining. If successful, CRONOS will provide a blue print for other surveys wishing to develop cross-national probability-based web panels and generate further high quality comparative survey data.

Finally, online capabilities will be exploited under SERISS to develop a harmonised, fast, high quality and cost-effective coding module for socioeconomic variables such as occupation and education, key information for understanding patterns of and inequalities in behaviour. The module will be underpinned by large multilingual databases with tens of thousands of entries about job titles, industry names, fields of education and training, educational qualifications and employment status. By accessing this database, interviewers will be able to code all this information directly with the respondent during the interview process rather than having to rely on recording sufficient verbatim information to enable post-coding. Once developed, the module and associated databases could be used on any survey.

Reaching out

SERISS is committed to engaging with the scientific community across the social sciences and beyond to disseminate and exploit learning from the project. This includes liaising with the European Statistical System and national statistics institutes, which face many similar challenges. Stakeholder workshops on key topics such as new forms of data are planned, and we are seeking to establish a survey network to bring together data producers and users with different needs to learn from one another and from SERISS. Tools and software developed under the

Europe is fortunate to have a significant social science data infrastructure in place with social surveys and administrative data at its core

project will be open source so that they can be adapted and adopted by other infrastructures. SERISS is also developing a comprehensive package of face-to-face and online training around data management and analysis. Social science data is made widely available to the wider research community.

The improved working practices and innovations resulting from SERISS are intended to lead to improved trust and confidence in social science data, increased productivity of researchers and economies of scale resulting in higher quality outputs and cost savings.

Europe is fortunate to have a significant social science data infrastructure in place with social surveys and administrative data at its core. Social science microdata have contributed significantly to the understanding of individuals' attitudes, behaviour and experiences, and are widely used by both European and national policymakers. The infrastructures employ hundreds of researchers and interviewers, as well as interacting on a regular basis with the general public, i.e. survey data subjects. Three SERISS research infrastructures (ESS ERIC, SHARE ERIC and CESSDA AS) have been recognised as 'Landmarks' and another (GGP) as an 'Emerging Project' by the European Strategy Forum on Research Infrastructures (ESFRI).

Further work is necessary, however, if the social sciences are to remain fit for purpose and meet the demand for high quality data cost effectively, respond to the emergence of new technology and new forms of data, and engage with end users successfully. Continued investment in infrastructure building clusters such as that supported by SERISS is essential to facilitate this.

Dr Sarah Butt
Research Fellow
School of Arts and Social Sciences,
Department of Sociology
University of London

http://www.city.ac.uk/people/academics/
sarah-butt
http://seriss.eu/