



Research Infrastructures in Social Sciences and Humanities

Authors:

Maria Samara and Henry Scott

EKT NHRF

© 15 December 2014 NET4SOCIETY

www.net4society.eu

All rights reserved

Reproduction only with written consent by the coordinator

NET4SOCIETY is an FP7 project funded by the EUROPEAN COMMISSION

THEME 8: Socio-economic Sciences and Humanities

Coordination and Support Action

FP7-SSH-2012-2

SSH.2012.8.8-2

This publication reflects only the author's views

– the EU Commission is not liable for any use that may be made of the information contained therein.

Table of Contents

1.	Introduction	3
2.	Data from research in social sciences and humanities	5
3.	Research infrastructures in social sciences and humanities	6
4.	Research infrastructures in social sciences and humanities – the way forward	10

1. Introduction

Research in social sciences and humanities has created a large quantity of digital material and data with strategic importance for the European policies that address societal challenges at regional, national, European and world level. The results of these research activities are stored in institutional, national or European Research infrastructures, and are very useful for European policy-making and the scientific community, supporting multi-disciplinary and trans-disciplinary research, which is a priority for research in Horizon 2020.

All disciplines are affected by the endeavours towards information integration. The building & development of existing and new research infrastructures can help to integrate existing research communities and facilitate a broader exchange of results and advancing research carried out by the various disciplines. Research Infrastructures are crucial in developing the research capacities and strengths of Europe and in enhancing Europe as a knowledge-driven economy. Researchers need them to conduct effective and comparable research. In order to maximise the impact of research in all thematic areas, the services offered by Pan-European research infrastructures include:

- Improved access and data protection
- Improved integration of scientific collections
- Reliable and up-to-date information to conduct realistic studies and make credible recommendations
- Efficiency, productivity, and scientific quality across disciplines as a result of data comparability.

Research infrastructures exist in a surprisingly large variety of forms and structures and make an important contribution to the advancement of knowledge in all scientific areas. Research infrastructures are defined as "Facilities", "resources" and related "services" that are needed by the scientific community for the development of leading-edge research in the most efficient manner at EU level, as well as for knowledge transmission, knowledge exchanges and knowledge preservation. Research Infrastructures include major equipment, collections, archives and structured information, enabling ICT-based infrastructures (e-infrastructures), and any other entity of a unique nature used for research.

The achievement of the Lisbon goals is enhanced by the development of e-Infrastructures which help to build new research environments. The successful development of e-Infrastructures in sectors such as e-Government (especially procurement), e-Health etc., has been of real benefit to Europe. The 'virtualisation' of experiments enables researchers from all around the globe to cooperate and share data using advanced research networks, driving productivity and increasing the quality of science.

Under FP4 and FP5, funding from the European Commission for research infrastructures in SSH was primarily provided for access and networking activities of different individual projects such as library collections of the British Library of Political and Economic Sciences and the European University Institute Library, as well as to statistical databanks such as the Norwegian Social Science Data Services and the Central Archive for Empirical Social Research in Germany. The 6th Framework Programme has supported nearly 250 research infrastructures, but only around 15 projects in the social sciences and the humanities.

During the 6th Framework Programme, the European Commission has published a report including information about funded projects and their databases, indicators and tools, mainly addressing issues of the knowledge economy and society, welfare and demography.

Analysis of the results of the FP7 calls for proposals undertaken by the National Contact Point (NCP) network for Research Infrastructures (RI) demonstrated that SSH is under-represented in terms of the level of participation and success rate in the Research

Infrastructures Programme. There is potential for the SSH disciplines to engage more with the Programme, particularly in the area of e-Infrastructures. SSH researchers should be encouraged to engage with the Research Infrastructures programme since it offers many opportunities. This could be achieved through the collaboration of the RI NCPs with the Challenge 6 NCPs at national and European level through joint activities, as the latter can act as multipliers for the Research Infrastructures Programme. An example would be the high visibility of the publication “Opportunities for SSH researchers in Horizon 2020 calls” that also features funding opportunities in the Research Infrastructures part of Horizon 2020.

This report aims to provide evidence of the data produced from research in social sciences and humanities not only to enable SC6 NCPs to better advise their clients but also to bring to the attention of a growing number of interdisciplinary research fields the importance of the services offered by research infrastructures in the area of social sciences and humanities.

2. Data from research in social sciences and humanities

The *HERA Survey on Infrastructural Research Facilities and Practices for the Humanities in Europe report*¹ demonstrates that more than half of the nearly four hundred research infrastructures have been developed in the area of cultural heritage and a quarter concern the area of language and linguistics. Furthermore, more than 60 per cent of the reported instruments are databases such as catalogues, reference tools and digitised objects such as books, periodicals, sound and images and the majority of them are freely available on the internet. The most popular existing infrastructures reported are databases (catalogues, reference indexes, bibliographies, etc.) and the second most popular are digitised texts, objects, artefacts, sound, film, etc.

The FLASH-IT project² (Facilitating Access to Socio-economic Research through Information and Communications Technologies), funded by the 7th Framework Programme of the European Commission, conducted a survey addressed to all project coordinators in the Socioeconomic Sciences and the Humanities Programme (SSH) with the aim of assessing the current status on how data management is implemented in SSH funded projects. Respondents to this survey - researchers in sociology, history, urban sciences, public policy economic sciences - keep their project research data in Database files (e.g. MsAccess, MySQL, Oracle, DBASE), Image files (e.g. jpeg, tiff, gif, etc.), Spreadsheets (e.g. Excel/.xls), Text or text presentation formats (e.g. pdf, html, Word), Video (e.g. Video Quicktime, MPEG etc.) and store them mainly in their computer, server, institutional repository and/or external web service (e.g. Dropbox). Further to this, the majority of respondents reported that the data produced are only available to their research group (the project consortium and their institution) or are available on a fee-paying basis (e.g. toll publication, specialised database etc.). Data protection and other confidentiality issues are the main factors that discourage them from data sharing while most of them used a research infrastructure for conducting their research. All of the respondents support the importance of Research Infrastructures for preserving and accessing EU funded project data.

In 2011, midway through the 7th Framework Programme³, the European Commission published a report entitled "Databases from socioeconomic research projects for policymaking" which assessed the impact of research and innovation on competitiveness using the data that emerged from the socio-economic research funded under FP7. This report also presented, thus far, the socio-economic datasets often used to derive composite indicators: datasets relating to the internationalisation of European firms, global innovation networks, RTD indicators, fiscal and monetary policies, Input-Output tables, the performance of universities, entrepreneurship, job flows, and measures of economic growth, productivity and employment.

The report illustrated the fact that all disciplines are affected by the informational integration and the building and development of existing and new research infrastructures.

As a cross-cutting issue of broad relevance, research in Social Sciences and Humanities is integrated into all disciplines, and therefore knowledge and promotion of research infrastructures hosting SSH research outputs is highly important for National Contact Points in all thematic areas of Horizon 2020.

¹ http://heranet.info/system/files/HERAJRPdocuments/Deliverables/d7.1.1_hera_survey_on_humanities_infrastructures.pdf

² <http://www.flash-it.eu/project-overview>, SSH.2011.8.8-2; project number 290431

³ ftp://ftp.cordis.europa.eu/pub/fp7/ssh/docs/databases-from-ssh-research_en.pdf

3. Research infrastructures in social sciences and humanities

In 2002 the European Commission initiated the **European Strategy Forum on Research Infrastructures (ESFRI)** with the aim of strengthening the cooperation between research infrastructures at European level and thus increase the competitiveness of European Research. This forum includes representatives from member states and the European Commission and aims at developing a European Strategy for RIs. More specifically, four working groups, specialised in different thematic areas including social sciences and humanities, publish their plans, or 'roadmaps', that list infrastructures of Pan-European interest which are either in a preparatory (design), approval or implementation (construction) phase. The latest roadmap includes three social science infrastructures and two in humanities. These infrastructures are listed below:

ESFRI Research Infrastructures - Social Sciences

- **ESS - European Social Survey Upgrade**

Website: <http://www.europeansocialsurvey.org>, phase: implementation

- ESS is an academically driven long-term pan-European instrument designed to chart and explain the interaction between Europe's changing institutions and the attitudes, beliefs and behaviour patterns of its diverse populations. The original infrastructure was set up in 2001 as a time series survey for monitoring change in social values throughout Europe and to produce data relevant to academic debate, policy analysis and better governance, and as an important resource for training new researchers in comparative methods.
- The ESS ensures that academically rigorous, comparative data about the attitudes, behaviour and beliefs of Europeans are freely available. More than 60,000 people have registered online to use the ESS data.

The ESS questionnaires cover a wide range of topics – challenges for European society – as follows:

- Moral and social values
- Health and well-being
- Trust in institutions
- Education and occupation
- Social capital and social trust
- Household circumstances
- Citizen involvement and democracy
- Social exclusion
- Political values and engagement
- Socio-demographic characteristics
- Immigration
- Crime

More than 30 countries across Europe took part in the six rounds of data design, collection and dissemination of ESS.

Further to this, ESS offers training in the form of face-to-face meetings or online courses. The ESS training courses "ESS Train" are two-day courses

which focus on key aspects of the survey lifecycle from a comparative, cross-national perspective. The specific aim is to equip researchers with the skills and knowledge they need to improve the rigour and equivalence of cross-national survey research in the European context.

Additionally, ESS has contributed to improving the methodology of other surveys, for example, the European Quality of Life Survey.

- **CESSDA - Council of European Social Science Data Archives**

Website: <http://www.cessda.org> , phase: implementation

- CESSDA is a distributed RI that provides and facilitates access to high quality data for researchers and supports their use. It promotes the acquisition, archiving and distribution of electronic data, and encourages the exchange of data. The RI includes 20 social science data archives in 20 European countries. Collectively they serve over 30,000 researchers, providing access to more than 50,000 data collections per annum.

The thematic areas of the data catalogue are listed under the following link: <http://www.cessda.net/catalogue/>

- CESSDA was fully implemented in June 2013 when it was set up as a permanent legal entity known as the Consortium of European Social Science Data Archives (CESSDA AS). An analytical workplan is provided in the following link:

<http://www.cessda.net/export/sites/default/news/CESSDA-Workplan.pdf>

- **CLARIN - Common Language Resources and Technology Infrastructure**

Website: <http://www.clarin.eu> , phase: preparation

- CLARIN aims to provide the social sciences and humanities research community with easy and sustainable access to digital language resources (in written, spoken, video or multimodal form), and advanced tools to discover, explore, exploit, annotate, analyse or combine them, wherever they are located. CLARIN is building up a network of data repositories, service centres and centres of expertise that are accessible by all members of the academic community in all participating countries.
- Unfortunately, due to a lack of financial commitment by the majority of member states in supporting this infrastructure, it cannot yet proceed to the implementation phase. This would seem to indicate a lack of awareness of the infrastructure's socio-economic benefits at governmental level. Also, according to the recommendations of the high level expert group assessing the projects on the ESFRI roadmap, the commitment of the scientific communities other than that for linguistics needs to increase.

ESFRI Research Infrastructures - Humanities

- **DARIAH - Digital Research Infrastructure for the Arts and Humanities**

- Website: <http://www.dariah.eu>, phase: preparation
- DARIAH is a digital infrastructure for the study of source materials in cultural heritage institutions.

- DARIAH provides long-term access to and preservation of research data and digital heritage materials for the arts and humanities in Europe. DARIAH connects information users (researchers), information managers and information providers. It gives them a technical framework that enables enhanced data sharing among research communities.
 - The operational plan of DARIAH will be available during 2014. At the moment, DARIAH collaborates with various projects and initiatives with the aim of increasing the network of affiliated projects⁵.
 - The ARIADNE (Advanced Research Infrastructure for Archaeological Dataset Networking in Europe) infrastructure brings together and integrates existing archaeological research data infrastructures.
 - CENDARI (Collaborative European Digital Archive Infrastructure) involves a broad collaboration between historians, research infrastructures and technical experts.
 - CHARISMA (Cultural Heritage Advanced Research Infrastructures)
 - DASISH (Digital Services Infrastructure for Social Sciences and Humanities)
 - DiXiT (Digital Scholarly Editions Initial Training Network)
 - The European Holocaust Research Infrastructure (EHRI)
 - NeDiMAH (Network for Digital Methods in the Arts and Humanities)
 - The high level expert group assessing the projects on the ESFRI roadmap noted that DARIAH is not ready for implementation and recommended, among other things, that an analysis of the user community be carried out and possible synergies with other projects be identified.
- **SHARE - Upgrade of the Survey of Health, Ageing and Retirement in Europe**

Website: <http://www.share-project.org>, phase: implementation

- SHARE is a Data infrastructure for empirical economic and social science analysis of ongoing changes due to population ageing. It focuses on all key areas of life, including social and family networks, socio-economic situations and health. SHARE is coordinated centrally at the Mannheim Research Institute for the economics of Ageing and is accessible free of charge.
- SHARE is the upgrade to a lasting infrastructure of a multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks of more than 86,000 individuals aged 50 or over and their young partners. It is harmonised with the U.S. Health and Retirement Study (HRS) and the English Longitudinal Study of Ageing (ELSA) and has become a role model for several ageing surveys worldwide: KLoSA in South Korea, JSTAR in Japan, CHARLS in China, LASI in India and ELSI in Brasil.

Why SHARE is so important⁶?

Informing public policies on results of cross national comparative research within the framework of the fact that population ageing exerts pressure on European welfare system.

⁵ <http://dariah.eu/about/collaboration.html>

⁶ http://www.euroris-net.eu/sites/www.euroris-net.eu/files/success_stories/Share_fin_2.pdf

Who can use the data?

After filling out a user statement, every person with scientific affiliation can download the data for free as long as the data are used exclusively for scientific purposes.

Mapping of Research Infrastructures in FP7

The Network of National Contact Points for Research Infrastructures (EuroRIsNET+)⁸, in collaboration with the MERIL project⁹ gathered, organised and provided access to information on projects funded by the 'Research Infrastructures' part of the Capacities Programme in FP7, per scientific area. More specifically, 95 projects were identified in the domain of Social Sciences and Humanities¹⁰, covering areas such as demography, growth and languages. The visualisation tools of the mapping of these infrastructures show that 7% of projects funded under 'Research Infrastructures' concern the area of Social Sciences and Humanities¹¹, while these infrastructures are widespread all over Europe.

The above developments make it clear that researchers from the humanities and social sciences need access to reliable and large datasets in order to conduct their research. Therefore NCPs supporting researchers in this area should provide them with information about the services offered by ESFRI and other research infrastructures they can use for carrying out reliable research.

⁸ EuroRIs-Net+ (2011-2013) – the Network of National Contact Points for Research Infrastructures (RIs NCPs) – was the EU-funded project that provided guidance, practical information and assistance on Research Infrastructures in Europe. EuroRIs-Net+ builds on the work of its predecessor, EuroRIs-Net (2007-2011) which provided value-added services to support transnational cooperation among NCPs. This support facilitated a more effective absorption of EU RI funding, resulting in a more prominent promotion of the opportunities offered by Research Infrastructures - at the European and international level - and their impact on e-science.

⁹ <http://observatory.euroris-net.eu/euroris/facilities/subview?q=Social%20Sciences%20and%20Humanities&t=SCIENTIFICDOMAIN>

¹⁰ <http://observatory.euroris-net.eu/euroris/facilities/subview?q=Social%20Sciences%20and%20Humanities&t=SCIENTIFICDOMAIN>

¹¹ <http://observatory.euroris-net.eu/euroris/visualizations/acts>

4. Research infrastructures in social sciences and humanities – the way forward

Horizon 2020 – the EU research funding programme for the period 2014-2020 – addresses societal challenges such as an ageing population, food security, and energy efficiency. Research in the area of social sciences and humanities is integrated into the Work Programmes of these challenges. Therefore NCPs in all challenges, but mainly NCPs supporting Societal Challenge 6, should encourage researchers to make use of the services offered by these infrastructures, as this adds value to their results and increases the impact of their research.

As regards the ESFRI roadmap, the next updates are due to be announced in 2016¹² and will include information on the phase update of infrastructures mentioned in previous roadmaps as well as new infrastructures, where the expectation is that SSH ESFRI infrastructures currently in the preparatory phase will be constructed. Also, new infrastructures may be introduced.

The increasing amount of social sciences and humanities research and the challenge of sustainability in project outcomes – particularly at European level and beyond – generated with European-level support provides the main impetus for encouraging researchers to use existing research infrastructures. To this end, the European Commission requires that projects funded by the Horizon 2020 programme submit a data management plan at the proposal submission phase.

One of the most important challenges is the growing interdependence among scientific disciplines. There is an increasing trend of ‘cross-fertilization’ of data within the SSH community to incorporate and take advantage of data from other scientific disciplines. Indeed, all research disciplines are faced with great challenges with respect to data creation, management, curation and access.

Furthermore, the availability of complete, harmonized and innovative data is a pre-requisite for evidence-based European policies¹³. It plays an important up-stream role in evaluating the relevance, costs and benefits of new initiatives on, for example, the challenges of globalisation, the impact of an ageing population and the Community immigration policy. The ESRC (the Economic and Social Research Council) has published a case study for the impact of the use of the ESS (European Social Survey) data on research and policy in the UK¹⁴ which highlighted the importance of increased data access for better research outputs and evidence-based policy making.

Additionally, the European Commission has produced a Report entitled “Evaluation of pertinence and impact”¹⁵ which provides examples and indicators for measuring the impact of research infrastructures on science communities, research policies and the economy/industry, and support the monitoring and development of impact measurement.

¹² http://ec.europa.eu/research/infrastructures/pdf/20140909-143726_September%202009_Launch_Call_Roadmap_2016.pdf#view=fit&pagemode=none

¹³ http://ec.europa.eu/research/infrastructures/pdf/workshop_june_2011/11_esrf_habfast_paper.pdf#view=fit&pagemode=none

¹⁴ http://www.esrc.ac.uk/images/WERS_ESS_tcm8-26047.pdf

¹⁵ <http://ec.europa.eu/research/infrastructures/pdf/csri.pdf>

One of the concluding remarks is that there is evidence of a possible future realisation of the positive impacts on the wider society but there is little demonstration of actual impacts systematically being achieved beyond a few ad hoc examples of direct and indirect societal impacts regarding the use of research outputs. A few projects have realised some liaison with local communities, which is important but ultimately not enough to ensure wider societal impacts from government investment in research infrastructures.

Therefore, National Contact Points should collaborate and support the promotion of such case studies and good practices concerning the impact that the use of Research Infrastructures has on research and on society as a whole.