



Providing open access to survey data: CESSDA solutions for ethical concerns

Sami Borg

Director, Finnish Social Science Data Archive

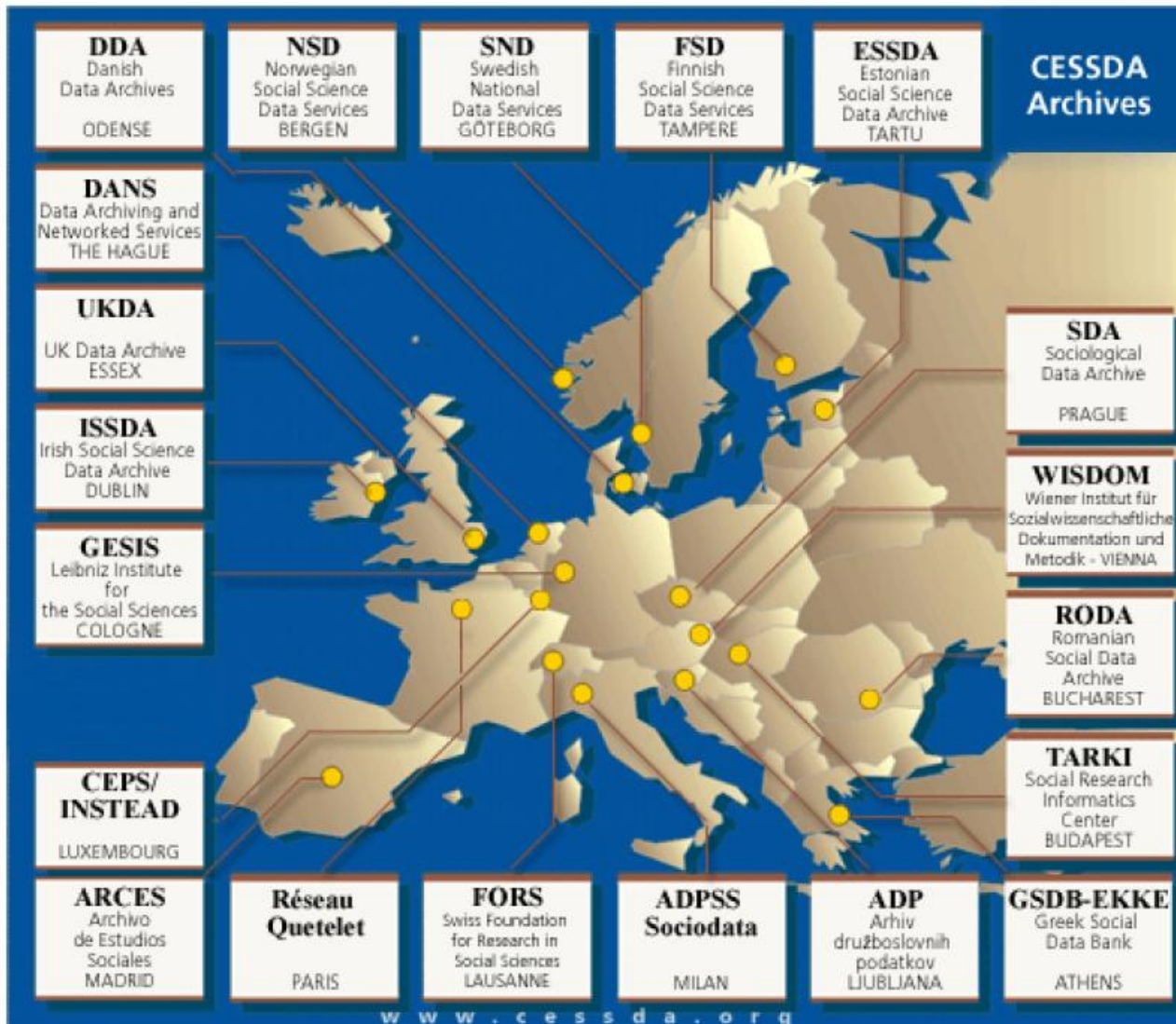
*Networking Event for European Research Infrastructures + Standards Workshop,
October 2, 2009, University of Helsinki, Finland*

*(This presentation is to great extent based on a document 'REFORMING THE COUNCIL OF
EUROPEAN SOCIAL SCIENCE DATA ARCHIVES' provided by the CESSDA PPP project)*



CESSDA

- *Council of European Social Science Data Archives was established in 1976, and it has currently 20+ members.*
- *Today CESSDA is a federation of national data dissemination and support organisations (data archives) spread across Europe.*
- *CESSDA holds General meetings and Expert seminars on a yearly basis*
- *It has a small, voluntary elected distributed executive.*
- *It is not formally a legal body, it has no central resources and exists on minimal funds.*
- *The member organizations are heterogeneous by their size. Most of them work in a university context, and have permanent funding for their national services.*





CESSDA...

- *Through data exchange agreements CESSDA provides access gateways to important social science data materials and to national collections generated across the public (government), academic and commercial sections, including: census collections; household, health and labour market data surveys; election studies; political opinion and social attitude surveys.*
- *It also serves as the gateway to important EC investments such as the European Social Survey, the Eurobarometers, the International Social Survey Programme and the European Values Surveys.*
- *In recent years CESSDA has developed resource discovery and data management and access tools, including Nesstar and multilingual resource discovery metadata tools.*
- *In 2008 CESSDA members held some 30,000 data collections, delivered over 100,000 datasets to more than 7,000 individual users and provided workshops and online information and services to some 200,000 researchers, learners and policy makers across Europe.*

Ethical concerns on survey data

- *Answers to ethical questions can be found from best practices*
- *Some guiding principles for such practices are listed in the OECD Principles and Guidelines for Access to Research Data from Public Funding (2007), and they do apply to social science survey data as well*
- *Survey information on human beings should be based on informed consent (IC). Whenever possible, the IC should allow secondary use of the data.*
- *Balancing between the need for anonymisation and trust in researchers (direct/indirect identifiers and the data disclosure risk).*
- *Social science data need documentation and contextual information, and it should be shared soon after (or even simultaneously with) the original research project.*
- *Many SSH research problems need fresh data, data might get old*

OECD Principles (see <http://www.oecd.org/dataoecd/9/61/38500813.pdf>)

- A. Openness*
- B. Flexibility*
- C. Transparency*
- D. Legal conformity*
- E. Protection of intellectual property*
- F. Formal responsibility*
- G. Professionalism*
- H. Interoperability*
- I. Quality*
- J. Security*
- K. Efficiency*
- L. Accountability*
- M. Sustainability*



CESSDA ERIC

- *The first ESFRI Roadmap was issued in 2006 which recommended 35 RIs across all disciplines, of which CESSDA was one. A revised Roadmap was issued in 2008, also including CESSDA.*
- *The CESSDA PPP project (2008-2010) aims to a new co-ordinating and administrative body recognised in international law and under-pinned by stable and long-term funding from national governments.*
- *It is intended for national funding bodies (e.g. ministries and research councils) and existing national data Resource Providers (e.g. data archives, current CESSDA member organisations).*



The aim of the new ERIC

- *A truly integrated European data infrastructure in which users can locate and gain access to data across the federation within a single integrated system, using a common set of protocols and procedures.*
- *In order to reach the aim...*
 - ...it is necessary to create a new legal body with a dedicated central body undertaking administrative duties across the network as well as developing common standards, technical platforms and middleware to be applied across the membership.*




The key benefit is...

Maximising the ability of social science and humanities (SSH) researchers and learners have access to the data resources necessary to conduct high quality research, teaching and learning, irrespective of the location of either individual or data within the European Research Area.

Other benefits



- *The existing CESSDA integrated data network will be expanded to include more resources both from within the existing framework as well as extending to more European countries;*
- *Coverage of the new research infrastructure will involve organisations beyond Europe;*
- *Cross-European resource discovery will be facilitated;*
- *Data which pose statistical disclosure risks are likely to become more readily available;*
- *The quality of data will be improved through the adoption and development of metadata standards;*
- *The certification of data archiving organisations will contribute to improved quality;*
- *Data archivists will have access to professional training;*
- *A European question bank of survey responses will become a real possibility.*



*In order to achieve these goals the
CESSDA ERIC independent legal entity
will work to facilitate:*

- *compliance with agreed interoperable metadata standards;*
- *adoption and application of a common single sign-on user authentication system;*
- *harvesting of all catalogue records by the data portal;*
- *downloadable data through common data gateways;*
- *maintenance of the 'local' language within the multi-lingual thesaurus;*
- *sharing of data archiving/preservation tools;*
- *adherence to OAIS reference model and approved certification of operational procedures and processes;*
- *contribution to cross-national data harmonisation;*
- *contribution to cross-national 'question bank';*
- *national agreement and adherence to 'open' access to publicly-funded research data, including where possible access to national government microdata.*

CESSDA ERIC will not be a data archive

- *Operating in a distributed fashion across a number of clusters of expertise, ideally it should neither hold data nor disseminate data.*
- *In terms of data discovery and access its role should be that of a gateway service – providing access (with common authentication) to a distributed*
- *set of national data collections.*
- *Its role will be as a co-ordinator, a facilitator and a broker.*
- *An additional key role will be in the development and implementation of standards and common tools, as well as ensuring where possible the adherence to 'good practice' through the provision of strategic guidance.*