

OECD GLOBAL SCIENCE FORUM

Changing Science: New Data for Understanding the Human Condition

Recommendations from the Expert Group on Data and Research Infrastructure for the Social Sciences

Presentation to the First DwB European Data Access Forum, Luxembourg, March 27th 2012

Peter Elias and Barbara Entwisle

Structure of the presentation

- Why was this Expert Group set up?
- What are its Terms of Reference?
- How has it functioned?
- What are the key recommendations it makes and how might these be enacted?
- What relevance does it have to a European Data Access Forum?



Why was the Expert group set up?

- Rapid growth in new forms of data collected in conjunction with commercial transactions, internet searches, networking, and the like.
- Technological advances in the capacity to access and link existing survey, census, and administrative data sets.
- Many developments happening worldwide to improve access to data for social scientific research
- Concerns about the unevenness of the developments and the problem of coordination between research funders, research communities and data producers
- Expert group to review developments in international data availability, consider their suitability for comparative research, detail the challenges to be addressed, and make recommendations to respond to these new opportunities.



Terms of Reference

ToR #1: To review and advise on the major data series which ideally would be available for research from every country, or initially at least from OECD countries,...which can be collected or combined to standardised formats...and made accessible for truly international comparative purposes.

THE UNIVERSITY OF WARWICK E-S-R-C ECONOMIC & SOCIAL RESEARCH COUNCIL THE UNIVERSITY OF JOHANNES GABRIEL WILHELM VON CEARAL, 1822

Terms of Reference

ToR #2: To review the potential availability for research over the next decade or longer of new forms of data generated by cyber and related activity...and how access to these for scientific...and policy research can be established.

THE UNIVERSITY OF WARWICK E-S-R-C ECONOMIC & SOCIAL RESEARCH COUNCIL THE UNIVERSITY OF JOHANNES GABRIEL WILHELM VON CEARAL, 1822

Terms of Reference

ToR #3: To review new developments and methodology for access to and digging into data in all its new forms, including disclosure analysis around confidential data, and the ethical issues involved in this.

THE UNIVERSITY OF WARWICK E-S-R-C ECONOMIC & SOCIAL RESEARCH COUNCIL THE UNIVERSITY OF JOHANNES GABRIEL WILHELM VON CEARAL, 1822

Membership of the Expert Group

Australia	Garth Bode	India	Dr S Durai Raju
Brazil	José Eduardo Cassiolato	Netherlands	Peer Scheepers Tom Srijders
Belgium	Patrick Deboosere Françoise Thys-Clément	New Zealand	Len Cook David Thoms
Canada	Chuck Humphrey	Norway	Bjørn Henrichsen
Denmark	Ole Gregersen Niels Ploug	South Africa	'Maseka Lesaana
European Commission	Pascal Jacques Philippe Keraudren Maria Theofilatou	United Kingdom	Vanessa Cuthill Peter Elias (<i>Vice-Chair</i>)
Finland	Sami Borg	United States	Barbara Entwisle (<i>Chair</i>) Myron P. Gutmann Julia Lane
France	Roxane Silberman	OECD	Stefan Michalowski Mika Shozaki
Germany	Gert G. Wagner York Sure Heinz-Herbert Noll		

Expert Group Meetings

Date	Location	Main topic
November 2-3, 2010	Paris	Project Initiation
January 24-25, 2011	Berlin	ToR #1
May 3-4, 2011	London	ToR #2
July 14-15, 2011	Venice	ToR #3
November 8-9, 2011	Washington, DC	Recommendations



Changing Science: New Data for Understanding the Human Condition

- Introduction and background
- A science driven approach to the global social science data agenda
- International data sharing: key problems
- What is already happening: work under way, plans in progress
- What needs to be done: recommendations
- How can these recommendations be taken forward?



**Introduction and background:
influential documents**

The slide features three document covers. The first, 'Riding the wave', is from The University of Warwick. The second, 'Harnessing the Power of Digital Data for Science and Society', is a report from the E-S-R-G Economic & Social Research Council. The third, 'Big data: The next frontier for innovation, competition, and productivity', is from The University of North Carolina at Chapel Hill.

Logos for THE UNIVERSITY OF WARWICK, E-S-R-G ECONOMIC & SOCIAL RESEARCH COUNCIL, and THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL are visible at the bottom.

Focus on digital data

Concerned exclusively with digital data in electronic formats:

- Census and survey records
- Administrative records
- Records of transactions, communication, internet activity
- Digital versions of books, newspapers, archived documents, etc.

Logos for THE UNIVERSITY OF WARWICK, E-S-R-G ECONOMIC & SOCIAL RESEARCH COUNCIL, and THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL are visible at the bottom.

**A science driven approach:
major global issues**

- population dynamics and societal change – migration, ageing populations, population growth, welfare/wellbeing;
- public health risks – spread of communicable diseases, lifestyle factors and non-communicable diseases;
- economic growth, innovation, research and development activity – science, education, workforce of the future, global trade and financial stability;
- social and environmental vulnerability and resilience– environmental change, dynamics of poverty and related policy evaluations.

Logos for THE UNIVERSITY OF WARWICK, E-S-R-G ECONOMIC & SOCIAL RESEARCH COUNCIL, and THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL are visible at the bottom.

**A science driven approach:
benefits from better access and sharing**

- Undertake comparative research to assess impact of national policies.
- Study rare groups, occurrences, or combinations of characteristics.
- National boundaries losing relevance as boundaries to human behaviour.
- Emergence of multinational commercial entities as data providers.
- Research collaborations extend across national boundaries
- Inference improved as number and diversity of countries increases.



**International data sharing:
discovering data**

Knowledge about what data are available and how useful these might be for specific purposes is an essential component of an international research agenda.



**International data sharing:
new forms of data**

- New forms of data (*e.g.* from the growth of electronic communication) have research value and will become ever more pervasive.
- Plans to determine their research potential will benefit from cross-nation coordination and collaboration.
- New data are likely to become increasingly important as the costs of traditional survey-based data collection continue to climb.



**International data sharing:
access issues**

- Nature of the consent(s) that data subjects may have given
- Need to protect and secure data that could reveal identities
- Problems of sharing across different legal jurisdictions

THE UNIVERSITY OF WARWICK E-S-R-G ECONOMIC & SOCIAL RESEARCH GROUP THE UNIVERSITY OF JOHANNES GABRIEL WILHELM VON CEASAR, 1822

**International data sharing:
problems of comparability**

- Essential that data definitions and classifications adhere to international standards.
- In some research areas, standards may still need to be developed further.
- Equally important is the availability and quality of associated metadata.

THE UNIVERSITY OF WARWICK E-S-R-G ECONOMIC & SOCIAL RESEARCH GROUP THE UNIVERSITY OF JOHANNES GABRIEL WILHELM VON CEASAR, 1822

What is already happening in the world of social science data?

- CESSDA (Council of European Social Science Data Archives)
- DwB (Data without Boundaries)
- DASISH (Data Service Infrastructure for the Social Science and Humanities)
- 'Digging into Data' (8 agency/3 country programme of research)
- Data Forums (Germany and the UK as examples) and PARIS21

Other organizations and institutions such as Eurostat (vision document and joint strategy for modernisation of the ESS), UNECE, the World Bank, WHO, ILO fund or undertake work to improve and/or expand data resources which are made available to the international social science research community.

THE UNIVERSITY OF WARWICK E-S-R-G ECONOMIC & SOCIAL RESEARCH GROUP THE UNIVERSITY OF JOHANNES GABRIEL WILHELM VON CEASAR, 1822

What is already happening in the world of social science data?

Linking Surveys to the World: Administrative Data, the Web, and Beyond

This project, administered by the US National Science Foundation and funded via the US Census Bureau will develop and evaluate methodologies that use the vast variety of data generated by households and businesses in the course of their ordinary activities. The project will examine administrative records created by businesses, individuals, and governments, streams of data from social media sites on the World Wide Web, and detailed geospatial data. The project will analyse these multiple source of data and relate them to data collected on surveys. It aims to improve survey measurements of economic and demographic data and potentially supplement or replace surveys with statistics based on administrative, Web-based, and geospatial data.

www.nsf.gov/awardsearch/showAward.do?AwardNumber=1131500&WT.z_pims_id=503587



What needs to be done?

Recommendations directed at:

- Research funding agencies, science ministries
- National Statistical Institutes
- International bodies with strong interests in socio-economic data
- The social science research community



Recommendations (1)

Exploring the research potential of new forms of data

Funding agencies should collaborate internationally to help specify and provide resources for a programme of research (methods and data types) that explores the research potential for new forms of data, their integration with traditional sources and linkage.

Official statistics and research users

Mechanisms should be found to bridge across the communities of official statisticians and social scientific researchers.



Recommendations (2)

Code of conduct covering the use of new forms of data for research

National research funding agencies should collaborate to develop a framework code of conduct covering the use of new forms of personal data, particularly those generated via network communication.

Improving incentives for international sharing of research data

International efforts be made to improve incentives for data sharing among scientific communities. Research funding agencies, publishers and employers of researchers to take coordinated actions to achieve this goal.



Recommendations (3)

Coordinating data management plans

Funding agencies should cooperate to share such information, publishing details of data management plans associated with new research awards.

International organisations and transnational data agencies

Leading international agencies (World Bank Group, World Health Organisation, United Nations Economic and Social Council, Organisation for Economic Cooperation and Development and the International Labour Office) should collaborate in the formulation of a strategic approach to the identification of obstacles to improved data sharing, the removal of these obstacles and a coordinated plan for the creation of data discovery tools on their websites.



Recommendations (4)

Global data curation

Social science research communities in countries without institutional support for data curation or supporting infrastructure should conduct an assessment of their national needs and assets in this area that will contribute to national plans of action. Working with researchers in such countries, established social science data archives should assist them by developing an assessment instrument and providing expert advice in preparing plans.



Taking the recommendations forward - monitoring and oversight

The Expert Group recommends that an oversight body should be established to monitor progress towards the implementation of these recommendations.

- This oversight body should have members appointed by participating countries.
- Participation of members would be funded by the organisations they represent.
- It would require a part-time secretariat, funded by national research funding agencies, to ensure that appropriate meetings could be arranged
- Two year pilot phase to take stock of willingness of countries to become engaged with recommendations



Relevance to the European Data Access Forum

Maximum advantage must be gained from activities already underway or planned. Where these have now or soon will have the relevant capacity, knowledge and expertise, these could provide appropriate platforms for some of the recommendations.

- DwB and EDAF
- NORFACE
- DASISH
- CESSDA
- DwB


