

Is the open data movement a threat or opportunity for NSIs?

Paola Tubaro

CNRS/Réseau Quetelet, Paris & University of Greenwich, London



Kiev, IAOS 2012 Conference, 12-14 September 2012



The “open data” context

- The “open data” movement is dramatically transforming official statistics and policy-making in many countries;
- Further opportunities ahead: expanding coverage to new areas, richer and more timely statistics.
- But integration of multiple sources and rapid publication raise issues of data quality and reliability of results;
- Expansion of open data is difficult when confidentiality protection is mandatory (detailed microdata);
- Potential competition from other data producers (private sector, web-mining...).

Open data and research data

- More detailed microdata are difficult to publish openly but access is usually authorised for scientific and statistical research purposes;
 - Established access systems at least in some countries, enabling collaboration between researchers and NSIs;
 - Yet gaps remain, unevenly distributed across countries: many recent efforts at improvements (in Europe: CESSDA PPP, DARA, DwB...);
 - In the open data era, pressure for scientific access is stronger than ever, and on the rise: yet another challenge for NSIs?
- While demanding greater attention, research access may provide new opportunities for NSIs.

Opportunities from research access

- Collaborative environments between researchers and statisticians, where they exist, may help to reap benefits from open data:
 - Developing standards for data and metadata;
 - Assessing data quality and educating policy-makers and the public about data quality issues;
 - Developing analytical tools for unconventional, less structured data;
 - Enhancing NSIs' use of data compiled by academic and other organizations in developing their own products.

A potential leading role for NSIs?

- NSIs that have experience allowing safe access to detailed data may set the example for other data-producing organisations;
- So far, this has been a grey area (recurrent “privacy crises” on Facebook, Google etc.);
- There is a strong need for guidelines on how to ensure confidentiality protection in general;
- NSIs’ experience may contribute to clarifying these issues, with great potential benefits for society at large.

A transnational approach

- This process could be most beneficial if countries where research access is most developed could transfer their know-how to those that have more limited experience
- Transnational collaboration would also improve research access *per se*: in Europe, access to multiple national datasets is currently difficult and fragmented
- Initiatives to improve research access at European level:
 - Data without Boundaries (DwB) involves NSIs, data archives and academics to pilot new solutions.

DwB in a nutshell

- **A four-year EC funded FP7-13 project (2011-2015)**
- **Linking the capacity of the research community with the important resources of government microdata in Europe**
 - **Coordination of existing infrastructures: CESSDA (European Data Archives) and the ESS (NSIs, NSAs, CB coordinated by Eurostat)**
 - **28 partners of which 1/3 are NSIs/NSAs, 1/3 Archives and 1/3 Universities**
 - **Also involving a wider range of stakeholders (European Data Access Forum)**

Conclusions

- **Enhancing research access as part of the overall debate on, and strategy towards, open data;**
- **Need to improve research access: transferring knowledge across countries, enhancing transnational solutions;**
- **How better research access may contribute to coping with “open data” demands:**
 - **Collaboratively developing tools for the management, analysis and dissemination of new data resources;**
 - **Raising awareness and spreading best practices to ensure greater data availability does not jeopardise confidentiality protection.**

Thanks for your attention

Website: <http://www.dwbproject.org/>

Contact: paola.tubaro@ens.fr



DwB
Data without Boundaries