Published in: ERCIM News vol. 2015 (2015) no. 100, pp. 19-20.

Open Data – What do Research Communities Really Think about it?

by Marie Sandberg, Rob Baxter, Damien Lecarpentier and Paweł Kamocki

Facilitating open access to research data is a principle endorsed by an increasing number of countries and international organizations, and one of the priorities flagged in the European Commission's Horizon 2020 funding framework [1][2]. But what do researchers themselves think about it? How do they perceive the increasing demand for open access and what are they doing about it? What problems do they face, and what sort of help are they looking for?

As a pan-European research data infrastructure, these are questions that are of fundamental interest to EUDAT. To better understand what researchers think, EUDAT has conducted a programme of interviews with fourteen major research communities from the fields of life sciences, Earth and atmospheric science, astrophysics, climate science, biodiversity, agricultural science, social science and humanities – a broad cross-section of European research interests. While the views of any given individual cannot be interpreted as the official position of a whole research community, they nevertheless provide useful information about the general attitude, requirements and challenges researchers face with regard to opening up their research data. In this article we report our initial conclusions from this survey.

Growing Awareness

Open access to research data is increasingly seen as a compelling principle in many research communities. There is a growing awareness of the global move towards open access, the potential benefits it may offer, and the need to implement open access policies within particular disciplines. According to preliminary figures on the first wave of open data pilot projects in Horizon 2020, the opt-out rate among proposals submitted to the "open by default" categories was below 30%, and the opt-in rate among other proposals was around about the same. This underlines our findings in EUDAT – researchers are pretty happy about sharing their data.

Challenges Ahead

In practice, though, there are many unsolved challenges still to be addressed, and those most often cited by researchers were the ethical and legal complications, and the issue of credit.

Not all data can be made open access. Personal data, and especially sensitive personal data, is particular challenging. In these days of large-scale combination and data mining, can such data truly be

anonymized for research purposes [3]? And what about the re-purposing of data for ends very far away from the original research agenda – for military or even criminal purposes? There are no easy answers to these questions, and the culture of ethics surrounding good research is making some communities tread warily.

Our survey highlights a lack of knowledge about the legal aspects of data sharing and data reuse, in particular around intellectual property rights, copyright and licensing, which can act as a barrier not only for opening data but also for re-using someone else's data. Choosing the right licence, for instance, can be a daunting task for some researchers who don't necessarily understand the implications of their actions.

While researchers are naturally keen to see their research published as widely as possible, in an interesting contrast to the open access scholarly paper movement, open data is viewed differently. Often research groups invest significant time and effort in collecting "hard to get data" which can then be used to build careers, offering what can only be termed a competitive advantage over those who do not have access to the same data. This issue of credit and consequent career progression is a real concern in many communities.

The way forward

While aware of, and supportive of, the open access data agenda, many research communities are looking for guidance about the practicalities of doing it; training on managing the legal issues, for instance. They also feel that these issues should be addressed at cross-disciplinary level, perhaps rendering the tasks even more challenging. And while much of the open access focus is on coordination efforts, training needs and policies, researchers also stress the importance of developing the right tools and services to enable these policies and, ultimately, the sharing and reuse of data; this is seen as particular crucial for handling sensitive data.

Some final words

Compared to scholarly publications, open access to research data is both less developed and more difficult to implement. Although open access to research data has only just begun, the broad spectra of expectations on EUDAT and other initiatives show that research communities have the notion that open access to research data cannot be solved through isolated activities or actions; instead it needs to underpin the whole system, reaching from strategic planning and overall polices to the mindset and everyday practice of the individual researcher.

Link:
EUDAT – European Data project:
http://www.eudat.eu/

References:
[1] European Commission: "Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020".
[2] G8 Science Ministers Statement, 13 June 2013, available at https://www.gov.uk/government/news/g8-science-ministers-statement
[3] Article 29 Data Protection Working Party: "Opinion 05/2014 on Anonymisation Techniques", adopted on 10 April 2014.

Please contact:
Marie Sandberg
CSC, Finland
Tel: +358 9 457 2001
E-mail: marie.sandberg@csc.fi

Rob Baxter
University of Edinburgh, UK
Tel: +44 131 651 3579
E-mail: r.baxter@epcc.ed.ac.uk