Policy recommendations for Open Access to research data in Europe – Stakeholder values and ecosystems

Thordis Sveinsdottir [a], Bridgette A. Wessels [a], Rod Smallwood [a], Peter Linde [b,*], Vasso Kala [c], Victoria Tsoukala [c] and Jeroen Sondervan [d]

[a] *University of Sheffield, Sheffield, UK*
[b] *Blekinge Institute of Technology, Karlskrona, Sweden*
[c] *National Documentation Center, Athens, Greece*
[d] *Amsterdam University Press, Amsterdam, The Netherlands*

Abstract. RECODE will leverage existing networks, communities and projects to address challenges within the open access and data dissemination and preservation sector, and produce policy recommendations for open access to research data based on existing good practice. The open access to research data sector includes several different networks, initiatives, projects and communities that are fragmented by discipline, geography, stakeholder category (publishers, academics, repositories, etc.) as well as other boundaries. Many of these organisations are already addressing key barriers to open access to research data, such as stakeholder fragmentation, technical and infrastructural issues, ethical and legal issues, and state and institutional policy fragmentation. However, these organisations are often working in isolation or with limited contact with one another. RECODE will provide a space for European stakeholders interested in open access to research data to work together to provide common solutions for these issues. RECODE will culminate in a series of over-arching policy recommendations for a policy framework to support open access to European research data targeted at different stakeholders and policy-makers (http://www.recodeproject.eu).

Keywords: Open data, Open Access, open research data, policies

1. Introduction

The objectives of the first RECODE work package are to:

- Identify and map the diverse range of stakeholder values in Open Access data and data dissemination and preservation.
- Map stakeholder values on to research ecosystems using case studies from different disciplinary perspectives.
- Conduct a workshop to evaluate and identify good practice in addressing conflicting value chains and stakeholder fragmentation.

Three related actions were used to address the objectives:

*Corresponding author: Peter Linde, Blekinge Institute of Technology, 37179 Karlskrona, Sweden. E-mail: Peter.linde@bth.se.

- An analysis of policy and related documents and protocols, in order to map the formal expression of values and motivations.
- Five case studies in particle physics, health sciences, bioengineering, environmental research and archaeology. These explored issues of data size; quality control, ethics and data security; replication of large datasets; interoperability; and the preservation of diverse types of data.
- A validation and dissemination workshop that sought to better understand how to match policies with stakeholder drivers and motivations to increase their effectiveness in promoting Open Access to research data.

2. The definitions of and vision for Open Access

- The European Commission definition of "Open Access" is "free... access to and use of publicly-funded scientific publications and data".
- The Berlin Declaration states that Open Access contributions include original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material.
- The Berlin Declaration's vision is that Open Access to data has the potential to create "a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community".

3. The stakeholder taxonomy

We have identified five basic functions in the Open Access ecosystem: Funders & Initiators, Creators, Disseminators, Curators and Users. These functions are represented by different performers (stakeholders). Each performer undertakes activities and produces records in relation to Open Access Data. The functional taxonomy of Open Access stakeholders was constructed in parallel to the review work and mapping of stakeholder values in WP1. The construction work of the taxonomy and the wide search for data management protocols, Open Access policies, and ethical protocols have fertilised and re-fertilised both tasks to ensure that we cover all aspects of the Open Access ecosystem.

4. Conclusions

There is a clear overall drive toward Open Data Access within the policy documents, which is part of a wider drive for open science in general. This is underpinned by the view of science as an open enterprise. In societal terms, data is valued as a public good in the sense that its production is funded by public money and thus should be accessible to the general public.

Discussions of Open Data tend to refer to science as a single sector, leading to differences between disciplines being ignored in policy making. Each discipline has different methods for gathering and analysing data, some disciplines deal with sensitive data, others deal with data that may have IPR or legal issues. It is important that these differences are recognised, as they will inform the debate about subject specific requirements and common infrastructures for Open Data Access.

5. Future work

Work package 2 – Infrastructure and technology will focus on issues surrounding open access and data dissemination and preservation infrastructure and technologies, including issues such as standards, interoperability, metadata, etc.

D2.1 – Infrastructural and technological challenges and potential solutions (*submitted March 2014*).

Work package 3 – Legal and ethical issues in open access and data dissemination and preservation will identify legal and ethical obstacles, barriers and solutions in relation to open access and data dissemination and preservation across Member States and third countries from the perspectives of a range of different stakeholders.

D3 – Legal and ethical barriers and good practice solutions (*submitted April 2014*).

Work package 4 – Institutional evaluation and support for open access data will focus on institutional practices and barriers in open access and data dissemination and preservation, including examining measures for evaluating data quality, integrity, impact and trustworthiness of data as well as barriers such as training, funding and infrastructure.

D4 – Institutional barriers and good practice solutions (*expected June 2014*).

Work package 5 – Policy guidelines for open access and data dissemination and preservation will consolidate the information from WPs 2–4 and review relevant open access and data dissemination and preservation policies at the European and Member State level and in third countries. It will identify policy gaps where the grand challenges discussed in WPs 2–4 are not being addressed and consolidate a series of policy recommendations.

D5 – Draft guidelines for different stakeholder groups on supporting open access to and preservation of research data (*expected September 2014*).

Work package 6 – Stakeholder engagement and mobilisation will create a taxonomy of open access stakeholders and consider how open access and data dissemination and preservation stakeholders might be best mobilised to implement the RECODE policy recommendations and maintain collaboration activities between different types of stakeholders.

D6 – Using existing open access networks to support policy harmonisation across Europe (*expected January 2015*).