Research Data Policy of the Alliance of German Research Organisations - Implications and Implementation

Hans Pfeiffenberger

Alfred-Wegener-Institute for Polar and Marine Research, Helmholtz Association

Helmholtz OA Workshop Energy, KIT, 2011-05-25, Karlsruhe
Cologne, Germany, 2009; Historical Archive drops into tube construction site

- This is what must not happen to the building of scientific knowledge!
- North-Rhine-Westphalia
- 10 Mio population
- 2 civil engineers for oversight of building
On The Shoulders of Giants

- Isaac **Newton** famously remarked in a letter to his rival Robert **Hooke** dated February 5, 1676 that:

„What **Descartes** did was a good step. **You** have added much several ways, and especially in taking the colours of thin plates into philosophical consideration. **If I have seen a little further it is by standing on the shoulders of Giants.**“

en.wikipedia.org/wiki/Standing_on_the_shoulders_of_giants
Giants today are ... 

- probably more like ordinary humans ;-) 
- Authors of articles 
- Creators of datasets 
- **May datasets be less reliable than articles, books? Why treat them different?**
ESF/EuroHORCs Vision for ERA (2008):

“... permanent access to ... quality assured research data”

- **Aim:** Reuse & Reproduce
- **Quality Assessment**
- **Persistent (and Open) Access, Licensing**
- **Digital Longterm Preservation**
- **Data provided and described by researchers**
- **Basic and advanced data infrastructure, provided by ???**
Principles for the Handling of Research Data

- adopted by the Alliance of German Science Organisations on 24 June 2010

- Alexander von Humboldt Foundation
- German Academy of Sciences Leopoldina
- Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)
- German Academic Exchange Service (DAAD)
- Fraunhofer-Gesellschaft
- Helmholtz Association
- German Rectors' Conference (Hochschulrektorenkonferenz - HRK)
- Leibniz Association
- Max Planck Society
- Wissenschaftsrat (German Council of Science and Humanities)
Preamble

- **Quality-assured** research data
- **cornerstone of scientific knowledge**
- independent of the purpose for which they were originally obtained, can often serve as the basis for further research.
  - (re-use, not mainly reproduce!)
- **Preserving** research data over the long term and making them available
- **strategic task** to which science and the humanities, politics as well as other parts of society, must contribute.
- **objective of supporting the quality, productivity and competitiveness of science** and academia,
  - (keep in mind: not to build a data bureaucracy!)
Preservation and accessibility

- In accordance with important international organisations involved in funding and performing research [1],

- the Alliance supports the long-term preservation of, and the principle of open access to, data from publicly funded research.
  - This principle shall be balanced against the scientific and legal interests of researchers.
  - The protection of the personal data ... obligations to third parties ... have to be taken into account.
  - The principles of good scientific practice must also to be observed [2].

The scientific disciplines

- The **ways of and conditions for access to research data**
- must be **developed separately for the individual scientific disciplines**
  - taking into account the **methods of data acquisition**,  
  - the volume and **potential for integration** of the data,  
  - as well as its **practical usability**.  
  - the respective **life cycles and usage scenarios** of the data in the specific research fields
Scientific recognition

- The provision of research data for further use is a service which benefits the sciences and humanities in their entirety.
  - not: it is / can be considered scientific work

- The Alliance encourages the recognition and support of this additional costly and time-consuming effort.
  - Which kind of recognition?
  - Leave the details to the disciplines, research funders, evaluation?
  - Acknowledges additional cost/time => additional resources
Teaching and qualification

- For those involved in research, an appropriate range of training and support services for professional data management must be made available, meeting the specific requirements of the different disciplines.
Use of standards

- Proper use of research data requires that the data are documented and provided with appropriate metadata ...
  - This is much more essential and more difficult for data than for books...

- ... in a standardised manner. Observing subject-specific requirements, standards, metadata catalogues and registries are to be developed in such a way that interdisciplinary use is also possible.
  - Do we need to square the circle??
  - All „grand challenges faced by society“ are multidisciplinary!!
  - It is a reminder for (over-)specialized data repositories!!
Development of infrastructures

- Sustainable research data management imposes **a wide range of technical and organisational requirements.**
  - We are not talking about yet another portal (grid, db, ...)
- These requirements must be defined through the cooperation of researchers and information specialists.
  - Again: no self-serving bureaucracies ...
    (and the researchers have to experience the impact of their own requirements!)
- Infrastructures are to be developed according to these requirements and, if possible, **interoperably integrated in international and interdisciplinary networks from the start.**
  - Don‘t create islands!! A danger with project-funded narrowly focused, national infrastructures
Making this a reality is a more difficult task than it may seem. To collect, curate, preserve and make available ever-increasing amounts of scientific data, new types of infrastructures will be needed. The potential benefits are enormous but the same is true for the costs.

Neelie Kroes, VP European Commission
Thank you!