Data users, data producers

Stéphane Goldstein
stephane.goldstein@rin.ac.uk

Presentation to the
UK Environmental Observation Forum

UK-EOF Data Solutions Workshop
Garden Organic, Coventry
25 June 2009
What is the RIN?

- A small policy unit funded by the four HE Funding Councils, the seven Research Councils and the three National Libraries.
- Aims to enhance and broaden the understanding of how researchers in the UK create and use information resources and services.
- Supports the development of effective policies and practices for researchers, institutions, funders, information professionals and all others involved in the research information landscape.
What do we do?

RIN’s remit covers information in all forms, and how this relates to libraries, data archives, research funders, HEIs, publishers as well as researchers and all involved in the research process.

This encompasses work on research data:

*To Share or Not to Share* – researchers as data creator

Case studies in the life sciences

New study on the benefits of subject-based research data centres
Research data are important

- **Two reasons for making data available:**
  to make them a part of the scholarly record so that they can be validated and tested
  so that they can be re-used by others

- *To Share or Not to Share* was published in June 2008
  sets out to determine whether or not researchers do make their research data available, and the issues they face in doing so.
  covered the following areas: astronomy, chemical crystallography, classics, climate science, genomics, social and public health sciences, systems biology and the Rural Economy and Land Use programme.
Researchers add value to raw data (data cleansing, coding, derivation...)  
But these activities are usually specific to given projects; rarely take place with the aim of making it more usable to third parties  
Sound metadata standards exist, but their use by researchers is limited  
Raw model data have little value other than for creator; observational data have value in raw state  
Benefits from NERC’s investment in data centres
Climate science: motivation & constraints for publishing data

- For smaller projects, often little thought given to publishing datasets
- Many climate scientists feel that there are few rewards for publishing data
- No strong culture for data sharing or publication, although NERC centres provide a solid environment for deposit and curation of datasets
Key conclusions

- Lack of explicit career rewards attributed to data sharing
- Data needs to be recognised and rewarded as a legitimate research output in the same way that publications are
- Need to take into account of the complexity of data creation and preservation
- Needs to be active promotion of data publishing and re-use by research funders and institutions
- Needs to be a consorted effort made to increase the discoverability, accessibility and usability of datasets
- Need to consider what approaches are most appropriate in assessing quality of datasets
New work on data centres

- Investigate the benefits of effective sharing and curation of research data
- Examine the long-term usage and impact of data curated by established data centres.
- Lead to improving the data sharing environment and expanding the practice of data sharing in the UK
- Provide information to help data centres themselves improve their service to users.