Multi-Disciplinary Research Data Infrastructures

Results from a Roadmap Project

Jens Klump
German Research Centre for Geosciences
Acknowledgements

- Project “Framework for an Interdisciplinary Research Data Infrastructure”
- May 2011 – April 2013
- Partners from different research organisations.
  - Helmholtz Association
  - Max Planck Society
  - German Rectors' Conference (Universities)
  - Leibniz Association
  - World Data System
- Funded by DFG (LIS).
There is a whole spectrum of requirements and single solutions for dealing with research data.

The project set out to survey the status quo and chart a way forward for the development of research data infrastructures.

Special attention was paid to gap analysis.
Interdisciplinary research data infrastructures offer the potential for economies of scale.

- Are there services for research data management that are generic? (storage, identifiers, data groupware, ...)
- At what scale? (institutional, disciplinary, national, ...)
Data Life Cycle

- The most concepts for dealing with research data follow a cyclical model.
- The common data life cycle model pays attention to actors but not to roles and responsibilities.
- The common data life cycle model is not aligned to standards for data management, e.g. OAIS.
Managing the Data Continuum

- **Private Domain**
  - Transfer
  - Simple Metadata

- **Group Domain**
  - Transfer
  - Enriched Metadata

- **Persistent Domain**
  - Publication

- **Access Domain**
  -
Project “Research Data Infrastructures”
Project Structure

- WP 1: Status Quo (MPI Nijmegen)
- WP 2: Technology (ZIH Dresden)
- WP 3: Organisation (AIP Potsdam)
- WP 4: Costs (DKRZ Hamburg)
- WP 5: Community Building (SUB Göttingen)
- WP 6: Synthesis (GFZ Potsdam)
Project Results

... will be presented in this session.

- Private Domain: Challenges handling the data deluge from the perspective of the scientist.
- Group Domain: Virtual Research Environments.
- Persistent Domain: Cost structures and risk management.
Looking into the Future

- Information technology development is still very fast.
- IT service life cycles are 3-5 years; too fast for single source research data infrastructures.
- How can we cope with this rapid development?
- How can we deal with disruptive innovation?
- Long-term strategies?
Intermediate project results were presented on the project homepage [www.forschungsdaten.org](http://www.forschungsdaten.org).

Results from this project, and from related projects, will be compiled in a research data wiki replacing the old project homepage.

The wiki will be maintained by a group of authors and editors from several organisations.