Handling of research data in collaborative research centres: experiences from inter-disciplinary, large-scale, long-term research projects

Constanze Curdt ¹,², Dirk Hoffmeister ¹, Tanja Kramm ¹, Ulrich Lang ², Georg Bareth ¹
¹ Institute of Geography, ² Regional Computing Centre, University of Cologne, Germany

14th International Digital Curation Conference | Melbourne, Australia | February 5, 2019
RDM in Collaborative Research Centres / CRC Transregios

- **RDM in research projects**
  - **limited** project duration and funding
  - **data availability** during and after project duration (e.g. data re-use/exchange)
  - mostly **project-specific RDM solutions** available
  - service / infrastructure development according to **project and user requirements**

- **Collaborative Research Centre (CRC) / CRC Transregio (CRC/TR):**
  - **coordinated funding program** of the German Research Foundation (DFG)
  - **university-based** research institution (often various locations); established **for up to 12 years**
  - research field is **innovative, essential, challenging, multidisciplinary & long-term designed**

  ➢ since 2007 **“Information Infrastructure Project”** (INF-Project)

  “The main purpose [...] is the **management of relevant data** collected [...] with the aim of enabling **systematic and long-term use of such data**.” (DFG 2014)
RDM in the CRC/TRR32, CRC1211 & CRC/TRR228 – Project Overview

  ‘Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modelling, & Data Assimilation’
  – **Institutions**: Universities Aachen, Bonn and Cologne & Research Centre Jülich
  – **Disciplines**: geophysics, soil & plant science, geography, meteorology, mathematics, chemistry
  – **Study site**: Western Germany
  – **Goal**: research on exchange processes btw. soil, vegetation & atmosphere
  – **RDM system**: TR32DB online since February 2008 at www.tr32db.de
RDM in the CRC/TRR32, CRC1211 & CRC/TRR228 – Project Overview

- **CRC 1211** (CRC 1211, www.crc1211.de, funding since 2016)
  - ‘Earth - Evolution at the Dry Limit’
  - **Institutions:** Univ. Aachen, Bonn and Cologne & Research Centre Jülich
  - **Disciplines:** geology, soil & plant science, geography, meteorology, biology
  - **Study site:** Atacama desert, Chile
  - **Goal:** research on mutual evolutionary relationships between Earth surface processes & biota
  - **RDM system:** CRC1211DB online since December 2016 at www.crc1211db.uni-koeln.de

---

CRC/TRR32 (2007-2018)

Introduction – projects – demands – RDM services – conclusion
RDM in the CRC/TRR32, CRC1211 & CRC/TRR228 – Project Overview

- **CRC / Transregio 228** (CRC/TRR 228, www.crc228.de, funding since 2018)
  - ‘Future Rural Africa: Future-making and social-ecological transformation’
  - **Institutions**: Universities Bonn and Cologne, Berlin & Münster, BICC
  - **Disciplines**: economics, anthropology, soil & plant science, geography, virology
  - **Study site**: Kenya, Tanzania, Namibia
  - **Goal**: relationship between land use change, social-ecological transformation & future-making
  - **RDM system**: TRR228DB online since November 2018 at www.trr228db.uni-koeln.de

**CRC/TRR228** (since 2018-2021; possible until 2030)

**CRC 1211** (since 2016-2020; possible until 2028)

Introduction – projects – demands – RDM services – conclusion
Demands of RDM in the CRCs

- **Identification of demands**
  - check funding proposal, join project meetings
  - user surveys (every funding phase) and personal interviews

- **Summary of demands**
  - number of data
    - potential project duration of 12 years → approx. 45-70 project sections
  - data variety
    - heterogeneous data; discipline specific file formats; various file sizes (few kB to GB)
  - DFG recommendations
    - collaboration with library / computing centre, re-use of existing infrastructures, technologies & tools
    - data availability beyond project duration (e.g. data storage for 10 years)
Establishing RDM Services for the CRCs (and special services)

CRC / Transregio 32:
- special section for purchased climate data incl. visualization tool
- online WebGIS for data visualization

CRC 1211:
- special section for project weather station data; data provision and visualization tool
- online WebGIS for data visualization

CRC / Transregio 228:
- special exchange platform for stories (‘Transdisciplinary Diary’); in prep.
- workshop series on RDM topics
Project Databases for the CRCs – technical infrastructure

Available at: www.TR32DB.de, www.crc1211db.uni-koeln.de, www.trr228db.uni-koeln.de

WebGIS

DOI Application

Data Search & Download

Metadata Input

Database Structure

Standardized XML Files

CRC Website (publications, reports)

* hosted at Regional Computing Centre Cologne (RRZK)
Conclusion

- **Project RDM services and systems / project databases**
  - limited duration due to funding
  - preferable project-specific solutions available
    - cooperation with libraries / computing centers / projects → reuse existing infrastructure / services
    - system design in close cooperation with scientists

- **Research data**
  - heterogeneous, discipline specific data formats and data size
    - data documentation, interoperable metadata
    - data quality (not covered by the project)

- **Dealing with scientists**
  - early integration and cooperation
  - service development in compliance with user requirements
  - adequate training and support is essential; patience and understanding
Handling of research data in collaborative research centres: experiences from inter-disciplinary, large-scale, long-term research projects

Constanze Curdt ¹,², Dirk Hoffmeister ¹, Tanja Kramm ¹, Ulrich Lang ², Georg Bareth ¹
¹ Institute of Geography, ² Regional Computing Centre, University of Cologne, Germany

14th International Digital Curation Conference | Melbourne, Australia | February 5, 2019