



TERRECO Data Flow: Gathering, Archiving, Management and Provision of Long-term Data



University of Bayreuth

Gerhard Rambold, Marianne Ruidisch, Thomas Koehler, Stefan Holzheu, Oliver Archner, Björn Reineking and John Tenhunen



Contact: gerhard.rambold@uni-bayreuth.de

Problem Statement:

Within the framework of projects in environmental science and socio-economics (e.g., analyses of social-ecological-systems, a great abundance and diversity of data is generated. Such data belong to very different data domains, and it is a challenge to create a functional data management structure, considering the necessary data entry formats and interfaces, storage units and access tools that allow long-term use of the distributed data sets in information analysis. Recent communiques of project management agencies and research foundations, including the Deutsche Forschungsgemeinschaft, indicate the importance of providing data in a format for sustainable archiving and subsequent acquisition from specific institutional repositories.

Overall Goals:

We are designing and implementing hardware and software components of a general system for data flows from doctoral and postdoctoral investigators of TERRECO to a general archive. The archive will support cohort 3 modelling studies and will be transferred to other long-term repositories. Data flows consider information from automatic measurement devices, maps, distributions of species and questionnaires used in sociological studies. We standardize processing from data entry and transformation interfaces to the final holding structures. In this context, the advanced data storage and management platforms DWB (developed at SNSB, München) and BayEOS-Goat (developed at BayCEER, University of Bayreuth) form technical backbone of the information system.

Major Data Processing Steps and Overall Data Flow:

- 1 Data Pre-processing: a) Guided (data, metadata, media) file name standardization according to TERRECO conventions; b) Guided mapping of data to TERRECO name space
- 2 Data Submission: Submission on TERRECO data portal
- 3 Data Primary Archiving: a) Data transfer into TERRECO database (based on DWB database); b) Original data file archiving
- 4 Data Curation: Data standardization, quality improvement
- 5 Data Provision and Publication: TERRECO portal, various interfaces (based on BayEOS database)
- 6 Optionally: KNU data mirror
- 7 (Final) Long-Term Archiving: Selected data forwarded to DFG and Korean data archives in accordance with TERRECO data policy.

