Building the Core. The Process to Standardize a Core Data Set on Research Activities in Germany
Contents

- Brief Overview: Wissenschaftsrat (WR) / German Council of Science and Humanities
- Core Data Set: Motivation, Aspirations, and Principles
- The Project
The German Council of Science and Humanities (Wissenschaftsrat) provides advice to the German federal government and the state (Länder) governments on the structure and development of higher education and research.

Administrative agreement between the Federal and state governments on establishing a Council of Science and Humanities (05. September 1957; version from 01. January 2008).
The Council’s Organizational Structure

Chairman

Plenary Assembly

Scientific Commission
32 votes
24 scientists
8 Personalities from Public Life

Administrative Commission
32 votes
16 state governments (1 vote per state)
16 Federal Government

Head Office
5 units with staff of circa 80
about 45 of which have scientific backgrounds

Committees and Working Groups
- complemented by external experts
- prepare recommendations and reports
Core Data Set

Motivation

- At HEI and non-university research institutions (NURI): Increasing increasing costs incur due to expanding demands for research-related data (HEI autonomy, transparency, controlling)

- At systemic level: Collection of comparable data is cumbersome, often impossible.

- Main reason: lack of standards, exacerbated by federal structure in Germany.

- Ramifications:

  - some HEI consider boycott of data collection.

  - also: consensus that standardization is necessary.

- WR has experience from a pilot study on a national research rating.

- WR issued recommendation to initiate CDS project.
Core Data Set

Aspirations

- Possibility to establish system-wide basis for comparison (at subject or institutional level)
- Reduce costs for academic institutions
- Create incentives to limit data inquiries
- Improve data validity (multiple use → internal and external cross checks)

- Even before start signals for widespread interest in participation in the project.
Core Data Set

**Principles**

- CDS is limited to data on research activities, which can be
  - clearly defined and delimited, and
  - are relevant for majority of subjects
- CDS standardizes only small set of data (core)
  - research staff, promoting young researchers, grants and projects, awards, patents, publications
- Complemented with „extra-nuclear“ elements
  - editorships, transfer spin-offs, received royalties, promoting young researchers (2), grants and projects (2).
- **No central database**, but decentralized control over data at HEI and NURI.
Core Data Set

Selected further principles

_ CDS standardizes formats and definitions for data exchange.

_ Data for exchange will often be at aggregate level (e.g. subject), while local datasets should be organized at elementary (disaggregated) level.

_ CDS not concerned with interpretation and evaluation of the data.

_ Existing data should be importable through interfaces.

_ Wherever possible data is to be collected within administrative processes.

_ CDS creates no new reporting obligations.

_ All CDS data is compatible with CERIF.

_ Adherence to privacy / data protection standards controlled throughout standardization process.
The Project

Core Data Set as a Process

- WR recommendation as first stage.
- Second Stage: project to prepare CDS specification.
  - financed by Federal Ministry of Education and Research.
  - duration 18 months, start summer 2013.
  - small number of pilot HEIs to test CDS in process.
  - supervised by WR - Supervisory Board.
- Third Stage: Implementation in HEI and NURI.
  - regular updates of specification (5 year intervals).
Project Setup

Project Management
- Head
- Heads of PGs

Supervisory Board (nominated by WR)

Project Group: Bibliometrics
- topical expertise

Project Group: Definitions
- administrative expertise

Project Group: IT, interfaces
- IT expertise

Project Group: Classification, thesaurus
- PR
The Project

Aiming for acceptance and standardization

Multifaceted project organization aims at

- definitions
- data standards (CERIF)
- as well as creation of acceptance within HEI and NURI but also among professional societies.
Thank you

for your interest.
**Time-line**

**Beginn**
- M1
- M2
- M3
- M4
- M5
- M6
- M7
- M8
- M9
- M10
- M11
- M12
- M13
- M14
- M15
- M17
- M18

**Zwischenbericht**

**Ende**

**Vorbereitung**

- WR setzt Beirat ein

**Organisatorischer Vorlauf**

- WR setzt Beirat ein

**Aufbau Projektmanagement**

- Organisatorischer Vorlauf

**Bundesweite Informationsveranstaltungen**

- Anhörungen (4)
- Festlegung Inhalte (2)
- Standardisierung Publikationsformate STUFE I (5)*
- Import Schnittstellen STUFE I (2)
- STUFE II (6)
- Import Schnittstellen STUFE II (2)

**Abgleich mit CERIF Standard**

- Technische Machbarkeit Pilotstudie

**Konsultation (Anhörungen & schriftlich)** (8)

**Koordination mit Software-Anbietern (5+)**

**Konsolidierung/ Konkordanz**

**Zwischenbericht an den WR**

**Vorabveröffentlichung der Spezifikation durch WR**

**Überprüfung alle 5 Jahre**

**dauerhafte Begleitung**

**EuroCRIS Membership Meeting – German National Session, Section II | Bonn | 13 05 2013**