The Research Core Dataset: standardizing the management and reporting of research information in the German science system

euroCRIS membership meeting
Münster – 19 November 2019

Sophie Biesenbender
Structure

1. Preliminary remarks
2. Development of the „Research Core Dataset“
3. Structure and contents of the „Research Core Dataset“
4. Implementation of the „Research Core Dataset“
1.1 The German science system

• Collection and processing of research information particularly fragmented in the German science system
  – distributed legislative and regulatory authority for the governance of research institutions (Federal and State Governments)

• New public management (NPM) since the 1980s
  – growing autonomy of institutions
  – growth of third-party and performance-based funding
  – increasing relevance of rankings and need for transparency
1.1 The German science system (ctd.)

- Increasing need for reporting of research information (in Germany)
  - external: rankings, ratings, (output-oriented) reporting to ministries and funding organizations, Statistical Offices, Council of Science and Humanities, implementation of performance-based funding systems
  - internal: development of internal governance processes, evaluation, controlling, management, communication
1.2 Standardization of research information

- 2013: Recommendations for a research core dataset (WR)
- 2016: Recommendations for the specification of the RCD (WR)
- 2020: RCD evaluation (WR)

- 1980: Recommendations for the harmonization of research reporting of HEIs (WR)
- 2005: Pilot study for a discipline-based research rating (WR)
- 2013-15: Specification of the research core dataset
- 2016-20: Helpdesk to support the implementation process
2 RCD development process

• Recommendations of the German Council of Science and Humanities (2013)
• Specification project for standards development
  – Project period: 2013-2015
  – Coordinator: iFQ (predecessor of DZHW Berlin)
  – Structure: four expert groups, partners: pilot institutions, non-university research institutions, scientific societies
  – Development process:
    • Development of first draft of RCD specification by expert groups in coordination with partners (alpha release)
2 RCD development process (ctd.)

- Evaluation and modification of draft specification by advisory board (working group of the Science Council) (beta release)
- Publication of beta release and open invitation to comment
- Revision of specification according to the comments (first release 1.0)

- Recommendations of the German Council of Science and Humanities (2016)
  - implementation of RCD specification (by research institutions and report-requesting organisations)
  - establishment of a helpdesk to facilitate implementation process
3 RCD structure and contents

• Principles of the RCD
  – not a (central) dataset in the strict sense (but a standard specification for the decentralized reporting and processing of RI)
  – does not create new reporting obligations for research institutions but intended to be used for existing ones
  – development of a “flexible” specification that can be customized for different reporting purposes
  – different layers: core-shell structure (1. layer: mandatory information, 2. layer: optional information, 3. layer: areas for further development)
3 RCD structure and contents (ctd.)

• RCD specification contains standard definitions for RI in six areas
  – Employed staff
  – Young researchers (including structured doctoral programs)
  – Third-party funding and budget
  – Patents and spin-offs
  – Publications
  – Research infrastructures

• Aggregate vs. basic data

• Different parameters and aggregation levels (flexible use)
3  RCD structure and contents (ctd.)

• Employed staff & Young researchers
  – **Full-time equivalents & Number of staff**
    • **Parameters**: gender, staff category, form of financing, highest academic degree, type of contract (permanent vs. fixed-term), citizenship, age group, type of job (full-time, part-time)
    • **Aggregation levels**: organisational unit, discipline, research field
  – **Number of doctoral students**
    • **Parameters**: gender, citizenship, start of dissertation (year), supervisor, structured doctoral programme, cooperation with other research institutions, country of graduation (PhD qualification)
    • **Aggregation levels**: organisational unit, discipline, research field
3 RCD structure and contents (ctd.)

- Third-party funding
  - Third-party funding
    - Parameters: type of funding organisation
    - Aggregation levels: organisational unit, discipline, research field, research project
  - Third-party funded projects
    - Parameters: title, name of project coordinator, consortium leader, start of the project, end of the project, type of funding organisation, grant number, amount of funding
    - Aggregation levels: organisational unit, discipline, research field
3 RCD structure and contents (ctd.)

• Publications
  
  - Lists of publications
    • Parameters: publication type, document type, author, title, publisher, source, identifier, format, language, access rights, peer-reviewed, thesis, resource, funding organization, grant number
    
    • Aggregation levels: organisational unit, discipline, research field
3 RCD structure and contents (ctd.)

- Person
- Patent
- Publication
- Third-party funded project
- Doctoral programme
- Research infrastructure
3 RCD structure and contents (ctd.)

- Mapping with CERIF

<table>
<thead>
<tr>
<th>CERIF</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Entities</td>
<td>293</td>
<td>1814</td>
<td>665</td>
<td>76</td>
</tr>
<tr>
<td>Attributes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Link Entities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language Entities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RCD/KDSF</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Objects</td>
<td>7</td>
<td>56</td>
<td>48</td>
<td>77</td>
</tr>
<tr>
<td>Attributes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relationships</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specialisations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3 RCD structure and contents (ctd.)
4.1 RCD implementation and outlook

- (Voluntary) implementation of the RCD standard by research institutions
- (Voluntary) adaptation of (internal & external) information, reporting and evaluation requests to the RCD standard
- Helpdesk to support implementation process (Science Council 2016)
  - Project period: 2017-2020
  - Coordinator: DZHW Berlin
  - Type of project: service and research project financed by Federal and State Governments
  - Focus: advisory service, documentation of external requests, organisation of workshops and horizontal networking, concomitant research and suggestions for the future development of the RCD specification
4.2 Survey on RCD implementation

- Survey period: May-June 2019
- Return: representatives of 109 out of 195 public HEIs
  - 58 of 88 (66%) of universities (with the right to award doctoral degrees)
  - 51 of 107 (48%) of universities of applied science
- 34% of the HEIs plan to implement the RCD standard, in 66% of the HEIs no decision has been made so far
4.2 Survey on RCD implementation (ctd.)

- Preparation of RCD implementation
- RCD implementation started
- Do not know

Comparison between Universities of applied science and Universities.
4.2 Survey on RCD implementation (ctd.)

[Bar chart showing percentages for CRIS in use, CRIS in planning, and CRIS not in planning for two categories: No decision regarding RCD yet and Decision to implement RCD.]
Thank you very much for your attention!

Further information (in German, translation into English in progress):

http://www.kerndatensatz-forschung.de/

Dr. Sophie Biesenbender
German Centre for Higher Education Research and Science Studies
Schützenstraße 6a | 10117 Berlin | www.dzhw.eu | Germany
eMail: biesenbender@dzhw.eu