

Emergence of new standards for big data

Introduction

There are approximately 250.000 researchers (full-time equivalent) in Germany alone which corresponds to a worldwide percentage of 3-4%. Threshold countries are seeing double-digit growth rates and even today millions of scientific reports are being published every year.

Considering these facts, the establishment of new methods for automated data interpretation, analysis and documentation seems inevitable. One way to achieve this is usually the establishment of common (meta) data standards. Globally connected sciences lead the way and will quickly realize the benefit of improved data availability.

„Best practice standard“ in the social and economic sciences

Researchers need high quality data. A new german website titled auffindenzitieren-dokumentieren.de provides a concise overview over established data sources for the social and economic sciences and explains how researchers from those fields can practise good data management regarding those three topics (data location, citation and documentation). While this site is a great resource for scientists from those disciplines, it may also serve as an encouraging example for how practical knowledge can be beneficially accumulated in one place.

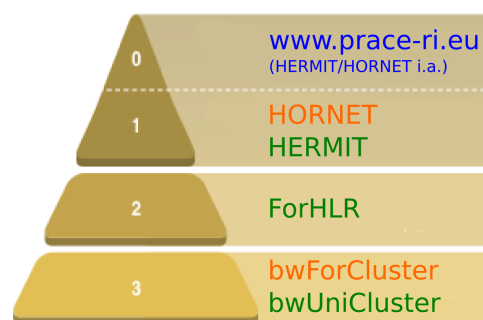
New Journal „Scientific Data“

The renowned Nature Publishing Group launched a new journal titled „Scientific Data“ this spring. The aim is to have a journal for high quality and detailed descriptions of the creation and content of an important publically available data set. Acceptance criteria for publication in the journal include high relevance and quality of the data and their descriptions. In-depth analysis and derived hypotheses are expressly not welcomed and belong in other journals. [1]

This way the generation of data (and making them accessible) will gain more importance as an independent academic contribution in accordance with long-standing demands by experts.

First tier of new high performance computer for Baden-Württemberg now operational

Scientists in germany may [apply](#) now for a [ForHLR I](#) account. The request will be appraised by the [steering committee of HLRS](#). Moreover, [Hornet](#) will be fully operational by December. The different discipline-specific bwForCluster will follow.



bwHPC-C5 service tiers from Baden-Württemberg to the European level. For access to the basic tier (3) only registration is needed. Access to the higher levels will be granted on a project basis.

Project info box

Interview status

The interview phase ends in November. More than 1200 user stories are currently being reviewed.

Requirements frequently identified by researchers include :

- clear-cut requirements and responsibilities (e.g. on the part of funding institutions like the DFG)
- better information about available RDM support, partners and opportunities
- common standards for formats, workflows and services
- easier processes (e.g. flexible "fair use" rules for scientists)
- miscellaneous services: digitization, access, processing, analysis and storage of data
- longer project funding periods
- more IT staff with general and discipline-specific expertise

<http://bwfdm.scc.kit.edu/>

<http://bwfdm.scc.kit.edu/mitarbeiter> (List of contact persons)

✉ Frank.Tristram@kit.edu (Editor)

[1] <http://www.nature.com/sdata/for-authors/submission-guidelines>