

## Aufkommen neuer Standards für „Big Data“

### Einleitung

Etwa 250.000 Wissenschaftler (Vollzeitäquivalent) gibt es in Deutschland, was einem weltweitem Anteil von 3-4% entspricht. Schwellenländer verzeichnen zweistellige Zuwachsraten und schon jetzt werden jedes Jahr Millionen von Arbeiten publiziert. So scheint es unausweichlich Methoden zu etablieren, die es ermöglichen neues Wissen automatisiert analysierbar aufzuschreiben. Wege dahin führen meist über gemeinsame (Meta)Datenstandards. Global verteilte Wissenschaften übernehmen hier eine Vorreiterrolle und sollten sich auch des Gewinns für ihr Fach durch bessere Datenverfügbarkeit bewusst sein.

### „Best Practice Standard“ in den Sozial- und Wirtschaftswissenschaften

Wissenschaftler benötigen hochwertige Daten. Die neue Webseite [auffinden-zitieren-dokumentieren.de](http://auffinden-zitieren-dokumentieren.de) stellt die gängigen Datenquellen für deutsche Sozial- und Wirtschaftswissenschaften in einer Handreichung zusammen und erklärt wie guter Umgang mit Daten für jeden dieser drei Themenbereiche praktiziert wird. Die Texte sind lesenswerte Wegweiser für Wissenschaftler dieser Disziplinen. Für andere Bereiche mögen sie als ermutigendes Beispiel dienen, wie Praxiswissen einer Fachdisziplin gewinnbringend zusammengetragen werden kann.

### Neues Journal „Scientific Data“

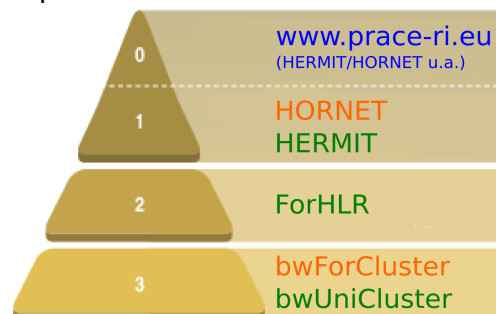
Die renommierte Nature Publishing Group hat im Frühjahr das neue Journal „Scientific Data“ herausgebracht. Sinn dortiger Publikation ist die (hochwertige) Erzeugung und den Inhalt eines öffentlich zur Verfügung gestellten Datensatzes im Detail zu beschreiben. Akzeptanzkriterien für das Journal sind die hohe Relevanz und Güte der Daten und deren Beschreibung. Tiefgehende Analysen und abgeleitete Hypothesen sind ausdrücklich unerwünscht und gehören in andere Journale. [1]

Damit gewinnt die Datenerzeugung (und Bereitstellung) langsam die Bedeutung einer eigenständigen wissenschaftlichen Leistung, was in Expertenkreisen auch regelmäßig so gefordert wird.

### Erste Stufe des neuen parallelen Hochleistungsrechners in Baden-Württemberg (ForHLR I) betriebsbereit

Forscher aus ganz Deutschland können einen [Projektantrag](#) zur Nutzung von ForHLR I stellen. Die Begutachtung der Anträge erfolgt durch den [HLRS-Lenkungsausschuss](#).

Auch [Hornet](#) wird bis Dezember vollen Betrieb aufnehmen. Später folgen die verschiedenen fachspezifischen bwForCluster.



Die Leistungsebenen des [bwHPC-C5](#) Konzeptes von der Landesebene bis Europa. Die Basisebene (3) bedarf nur der Anmeldung. Zugang zu höheren Ebenen wird zweckgebunden vergeben.

## Projektbox

### Stand unserer Interviews

Die Interviewphase endet im November, über 1200 User Stories sind im Reviewprozess.

### Beobachtete Themenfelder häufig genannter Wünsche von Forschern:

- Klarere Anforderungen und Verantwortlichkeiten (z.B. seitens Fördergeber)
- Bessere Information über FDM-Hilfen, Partner und Möglichkeiten
- Einheitliche Standards bei Formaten, Arbeitsweisen, Services
- Einfachere Abläufe (z.B. flexibles "Fair Use" für Wissenschaftler)
- Verschiedene Services:
  - Digitalisierung, Zugriff, Aufbereitung, Speicherung und Analyse von Daten
  - Längere Förderzeiträume zur Ermöglichung langfristiger Initiativen
  - Mehr Personal mit sowohl gutem IT als auch disziplinspezifischem Fachwissen

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

<http://bwfdm.scc.kit.edu/mitarbeiter> (Liste der Ansprechpartner)

[Frank.Tristram@kit.edu](mailto:Frank.Tristram@kit.edu) (Redaktion)

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