

Forschungsdatenmanagement

Ilvio Bruder



Warum Forschungsdatenmanagement und -archivierung?

WIRED

/SCIENCE

Ein Software-Bug könnte 20 Jahre Hirnforschung in Frage stellen

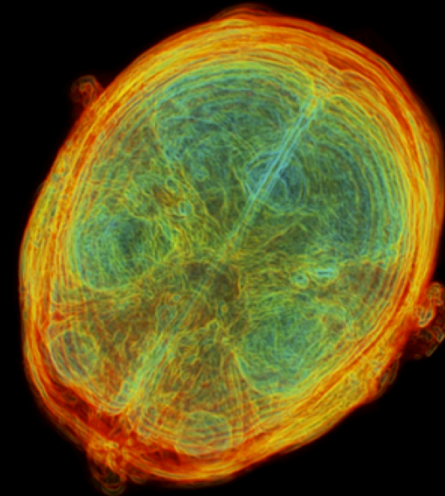
WIRED Editorial

05.07.2016

D

ie vergangenen 20 Jahre der Hirnforschung könnten gerade mit einer eleganten Handbewegung vom Tisch gewischt worden sein: Schwedische und britische

Medizininformatiker haben einen Bug in der Software entdeckt, mit der Magnetresonanztomographie-Aufnahmen analysiert werden.





Motivation: Warum Forschungsdatenmanagement?

- Forschungsdaten sind wertvolle Ressourcen
- FDM-Pläne in EU-, DFG- und BMBF-Projektanträgen gefordert
- Nationale Forschungsdateninfrastruktur für die kommenden 10 Jahre mit 80 Mio. € gefördert
- kollaboratives Arbeiten

Forschungsdaten (-Management) nach Kindling, Schirnbacher 2013

Forschungsdaten

"Unter digitalen Forschungsdaten verstehen wir alle digital vorliegenden Daten, die während des Forschungsprozesses entstehen oder ihr Ergebnis sind."

Forschungsdatenmanagement

"Der gesamte Prozess, der die Allokation, die Generierung, die Bearbeitung und Anreicherung, die Archivierung und Veröffentlichung von digitalen Forschungsdaten selbst oder von einer entsprechenden klassischen Textproduktion unterstützt, wird [...] unter dem Begriff Forschungsdatenmanagement zusammengefasst."

Open Research Data und FAIR-Principles

FAIR-Principles:

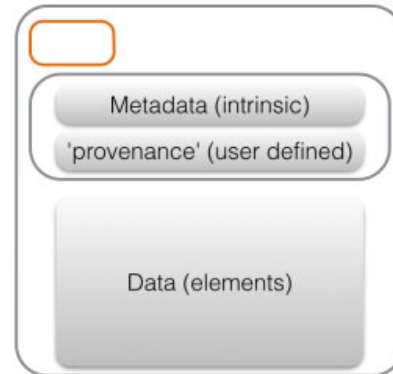
Findable

Accessible

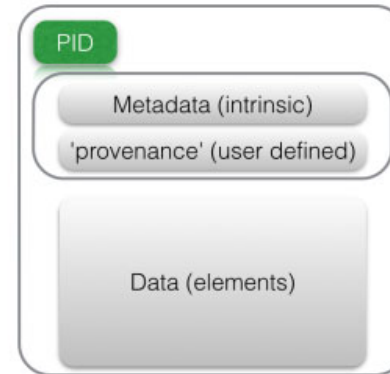
Interoperable

Re-usable

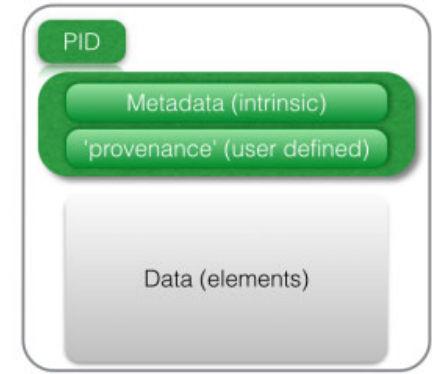
Totally UNFAIR



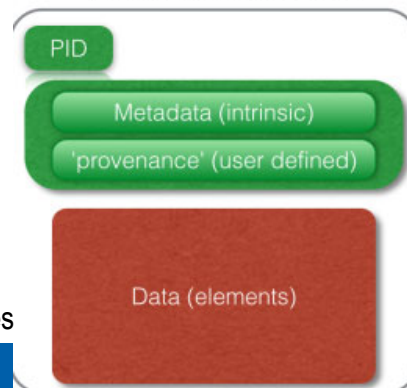
Findable
Usable for Humans



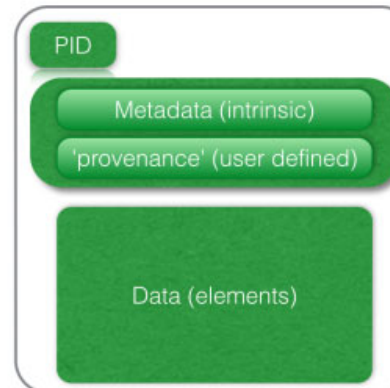
FAIR metadata



FAIR data-
restricted access



FAIR data-
Open Access



FAIR data-
Open Access/Functionally Linked



Forschungsdaten-Lifecycle



von fdm.uni-koeln.de/data-management.html

nach Ludwig, J & Enke, H 2013, Leitfaden zum Forschungsdaten-Management

Facetten des Forschungsdatenmanagement

Welche Daten benötige ich?

Wie sollen die Daten strukturiert werden?

Welche Werkzeuge benötige ich zur Verarbeitung?

Welche Personen haben Zugriff bzw. welche Verantwortlichkeiten existieren?

Müssen ethische oder rechtliche Aspekte beachtet werden?

Unter welchen Auflagen (Lizenzen) soll auf die Daten zugegriffen werden?

Benötige ich besonderen Schutz für meine Daten?

Wo möchte ich Daten veröffentlichen?

Wie lange sollen die Daten aufbewahrt werden?

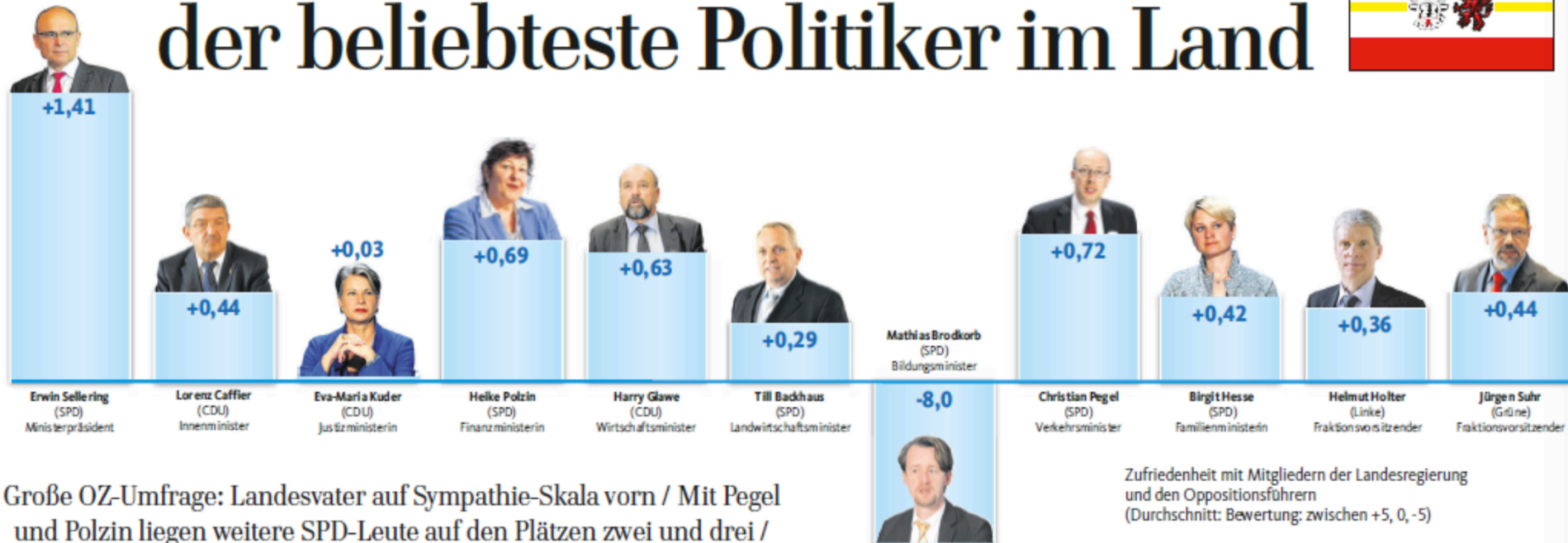
...

Forschungsdatenmanagement: Daten, Workflow und Provenance

- Welche Daten aufheben?
 - Primärforschungsdaten
 - Sekundärforschungsdaten
- Warum sollte der Forschungs-Workflow und die Methodik ebenfalls gespeichert werden?
 - Plausibilität der Ergebnisse
 - Nachvollziehbarkeit der Ergebnisse
 - Reproduzierbarkeit der Ergebnisse
- Wie greift man auf die Daten in 10 Jahren zu?
 - Archivierung der Daten plus aller nötigen Zugriffsinformationen
 - Digital Preservation Paradox

Beispiel: Plausibilität von Daten

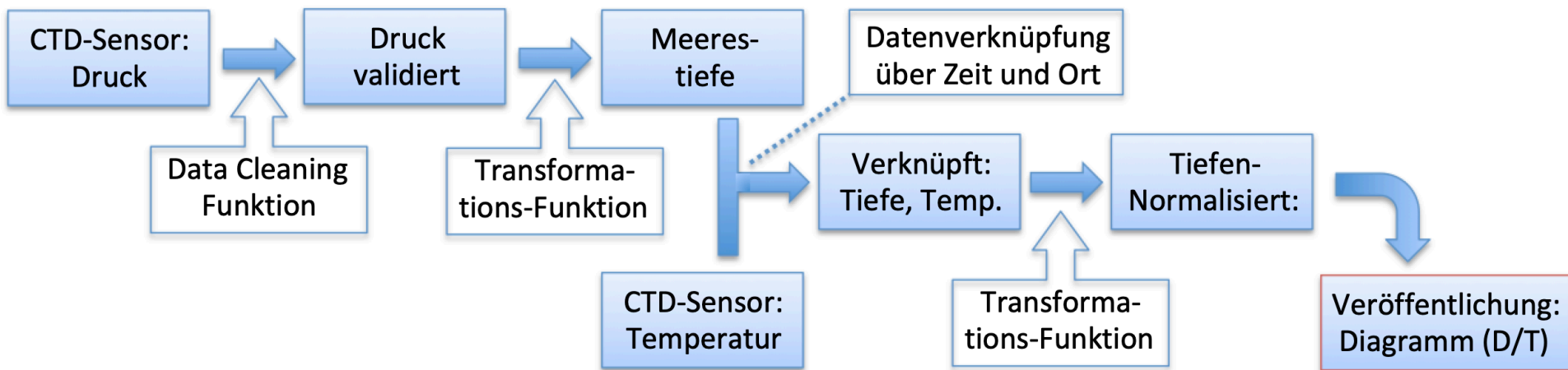
Ministerpräsident Sellering ist der beliebteste Politiker im Land



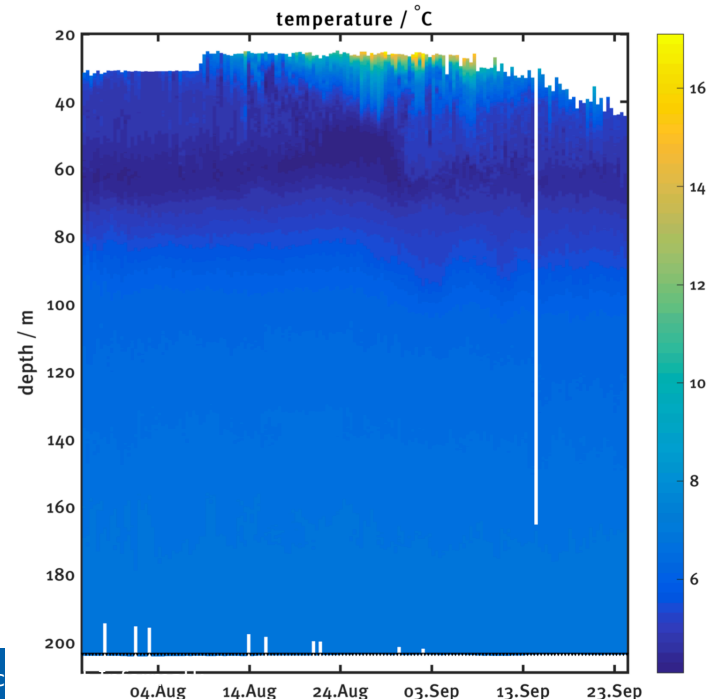
Große OZ-Umfrage: Landesvater auf Sympathie-Skala vorn / Mit Pegel und Polzin liegen weitere SPD-Leute auf den Plätzen zwei und drei / Gut 50 Prozent der Wähler sind mit der Landesregierung zufrieden

Zufriedenheit mit Mitgliedern der Landesregierung und den Oppositionsführern
(Durchschnitt: Bewertung: zwischen +5, 0, -5)

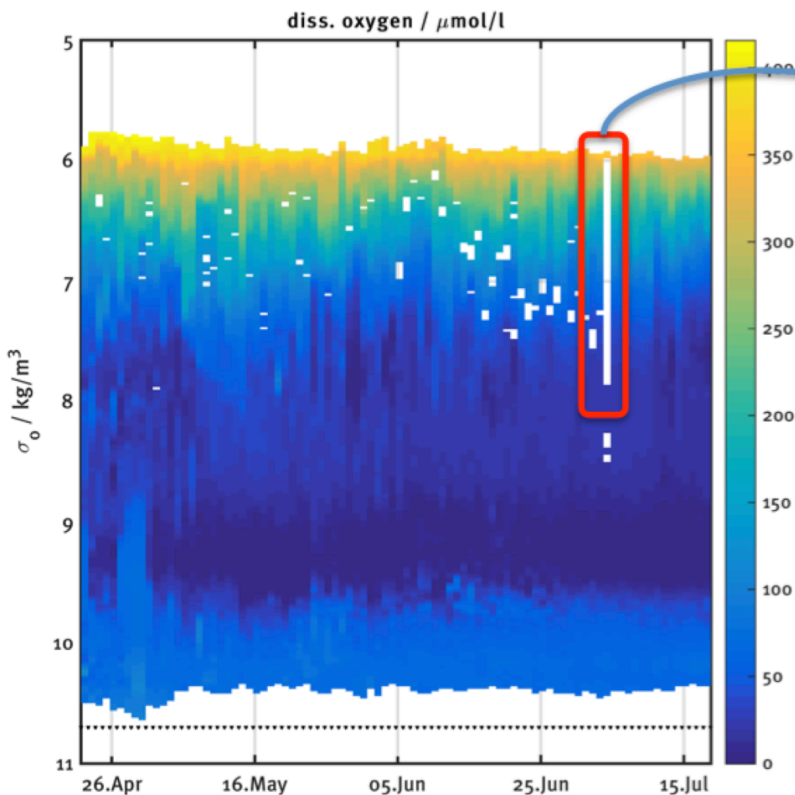
Konkreter Beispiel-Workflow am IOW



F	4PRES	PSAL	TEMP
S	7740	025	11 1
T	.5000000E+00	.1471000E+02	.5900000E+00
T	.1000000E+01	.1469000E+02	.6214348E+00
T	.2000000E+01	.1469000E+02	.6064539E+00
T	.3000000E+01	.1467607E+02	.6505998E+00
T	.4000000E+01	.1464342E+02	.6735354E+00
T	.5000000E+01	.1498016E+02	.6245019E+00
T	.6000000E+01	.1538000E+02	.6400000E+00
T	.7000000E+01	.1530750E+02	.7576394E+00
T	.8000000E+01	.1598441E+02	.8771477E+00
T	.9000000E+01	.1656642E+02	.9646380E+00



Beispiel IOW: Provenance-Anfrage



Why?
Provenance

Extensionale
Antwort

ID	Sauerstoff	Salz
...		
99		
0		
0		
0		
0		
...		
0		
101		

Intensionale
Antwort

Alle Werte des Sauerstoffgehalts
sind an dieser Stelle 0.
⇒ unwahrscheinlicher Wert
⇒ eventuell Sensorik ausgefallen



Werkzeuge an der Uni Rostock: Dokument-/ Projektverwaltung, Speichermöglichkeiten

- Sharepoint: Verwaltung von Dokumenten, Wiki, Kalender, Aufgaben, Diskussionen, Wiki
- Unibox: Clouddienst für Private Cloud
- ITMZ-Speicherlösungen: Virtuelle Server, Projektlaufwerk, Backup
- Gitlab: Versionskontrollwerkzeug, demnächst am ITMZ
- Tools und Informationen unter: <https://www.itmz.uni-rostock.de/onlinedienste/>

Werkzeuge an der Uni Rostock: Unterstützung für Planung, Verarbeitung und kollaboratives Arbeiten

- RDMO: Planungstool für Datenmanagementpläne
- Jupyternotebooks: Workbench zur Verarbeitung von Forschungsdaten
- CKAN: Ablage und Austausch von Forschungsdaten
- eLabFTW: elektronische Laborbücher
- RedMine: Projektmanagement-Tool

- zurzeit im INF-Projekt des SFB Elaine bereitgestellt und getestet
- z.T. auch durch nicht-projektbeteiligte Uni-Angehörige nutzbar
- Kontakt:
 - Frank Krüger, frank.krueger@uni-rostock.de
 - Max Schröder, max.schroeder@uni-rostock.de

Werkzeuge an der Uni Rostock: Unterstützung für die Datenpublikation

- Publikation in einem Fachrepository: z.B.
 - PANGAEA für Erd/Umweltwissenschaften
 - HEPData für Hochenergiephysik
 - DIP: Database of Interacting Proteins
- Uni Rostock Dokumentenserver: RosDok
 - DOI als referenzierbarer Identifikator
 - Nachweis auf RosDok sowie im K10plus-Katalog
- Allgemeine Repositories:
 - Figshare
 - Zenodo

Ros Dok

Rostocker Dokumentenserver

UB-Referat: Forschungsdatenmanagement

Sebastian Schick, Antje Meuser, Max Schröder, Ilvio Bruder

Aufgaben:

- Unterstützung bei der Publikation von Forschungsdaten (Datenpublikation auf RosDok)
- Beratung zum Forschungsdatenmanagement
- Unterstützung bei der Erstellung von Datenmanagementplänen
- Aktuell: Vorbereitung einer Forschungsdaten-Policy

Zusammenarbeit:

Prorektor für Forschung und Transfer, Servicezentrum Projekte: Projektakquise, ITMZ

Weitere Informationen:

Website: <https://www.ub.uni-rostock.de/wissenschaftliche-services/forschungsdaten/>

E-Mail: forschungsdaten@uni-rostock.de