



Die Matrix – funktionelle Anforderungen an Forschungsdatenrepositorien

RDA – Repository Platforms for Research Data
Ralph Müller-Pfefferkorn, TU Dresden, ZIH, Germany

research data sharing without barriers
rd-alliance.org

- Interest Group bei der Research Data Alliance
- Ziel: “... to improve the usability and technical capabilities of repository platforms.”
- Wie?
 - Zusammenbringen aller Akteure – Nutzer, Entwickler, Betreiber ...
 - Sammeln von Erfahrungen und Use Cases
 - Anforderungen definieren, Umsetzungen evaluieren ...
- Mögliche Ergebnisse: Empfehlungen, Wissenssammlung, neue Working Groups ...

- Gegründet 2014
- Co-Chairs
 - Robert Downs (Columbia University, USA)
 - Michael Jones (Elsevier, Australia)
 - Ralph Müller-Pfefferkorn (TU Dresden, Germany)
 - João Rocha da Silva (University of Porto, Portugal)
- „Ausgründung“
 - Working Group “Research Data Repository Interoperability”

- Sammlung von Use Cases von Repositorien für Forschungsdaten und deren Anforderungen
- Erstellung einer Matrix, die die Use Cases mit ihren funktionellen Anforderungen verbindet
- Start Herbst 2015

- Einfaches Use Case Dokument
- Zwei Teile
 - Kurze Beschreibung des Use Case
 - Tabelle mit Anforderungen

Requirement	Description	Motivation from Use Case	Importance (1 - very important to 5 - not at all important)

- 1. Scientific Motivation and Outcomes:** Provide a short summary of the scientific or technical motivation for the use case. What would be the best possible outcome and why?
- 2. Functional Description:** If possible, give one diagram/picture that indicates the overall structure/architecture you have/envision. Describe in simple words the functioning of the use case/system.
- 3. Achieved Results:** Describe results (if applicable) that have been achieved compared to the original motivation. What requirements could not be fulfilled and how did this influence the outcomes

Beispiel

Requirement	Description	Motivation from Use Case	Importance (1 - very important to 5 - not at all important)
Embargo date selection for data-depositing user	The user can select a specific date (e.g., 2016-05-15) or time span (e.g., 1 year) for when the research data shall become available.	This assure the user that their research data remains non-public for the chosen time period, and the university/funder that it will become available after that.	1
Assignment of DOIs	A DOI for the deposited data is generated upon deposit and resolves to the “landing page” for the research data, which displays only the required descriptive metadata during the embargo period	This reserves the DOI for the data, and makes the data citable with a DOI even before it's actually available	2

- Leider weniger Rückmeldungen als erwartet – nur 11 Einreichungen
 - ➔ Nicht repräsentativ
 - ➔ Aber interessante Einsichten aus der Praxis
- Aus verschiedenen Fachdisziplinen – Biologie, Medizin, Klima
- Am zahlreichsten generische Repositorien

- Erstellung einer Matrix
 - Zeile: Anforderung
 - Spalte: Use Case
 - Feld/Zelle: Bewertung der Anforderung
- Kombination überlappender Anforderungen
→ 43 Anforderungen
- Gruppierung in Kategorien
- „functional requirement score“ – Bewertung – je größer desto besser

- Authentication
- Data access
- Data and product quality
- Data organisation
- Integration
- Location
- Metadata
- PIDs
- Policy support
- Preservation and sustainability
- Publication
- Submission/ingest/management
- User experience/user interfaces

Die Matrix – Ausschnitt

Category	Requirement	Description	Functional requirement score	Number of Mentions	Dissertation	iRODS	DKRZ Climate
Metadata	Support for different metadata (schemas), including domain-specificity and interoperability	Any form of metadata (schemas) should be allowed. This includes metadata such as author, owner, license, source publication, librarian, date and time stamps; and domain-specific descriptors.	28	7		1	1
Persistent identifiers	Assignment of PID / DOI	At data ingest time or “project publication” or even before (e.g. when a paper was submitted but the data are not final yet) a PID has to be assigned to data and the collections it belongs to. It resolves to the “landing page” of the research data, which displays the required descriptive metadata during the embargo period. A clear transition for PID collections to a DOI has to be established.	26	7	2		1

- Matrix von funktionellen Anforderungen an Repositorien aus Use Cases erstellt
- Unterstützung beim Aufsetzen und Betrieb von Repositorien
- Hinweise/Leitlinie für Repositorien-Entwickler
- **Matrix** wurde Anfang 2017 als “Supporting Output” von RDA veröffentlicht

- Weitere Aktivitäten
 - Plenary 9 – Sammlung von Anforderungen aus Erfahrungen beim Aufbau von Repositorien (Ingest, Operations, Sustainability)
 - Plenary 10 – Reaktionen verschiedener Plattformentwicklern auf Anforderungen (Islandora, Fedora, Samvera, Dataverse, KIT DM, Dash)