



ZB MED – Leibniz-Informationszentrum Lebenswissenschaften

*Forschungsdaten in der Medizin - Management und Publikation
aus bibliothekarischer Sicht*



Relevante Aspekte des Forschungsdatenmanagements

- ▶ Was sind Forschungsdaten in der Medizin?
- ▶ Warum „Open Data“?
- ▶ Forschungsdatenmanagement und der Lebenszyklus von Forschungsdaten
- ▶ Rahmenbedingungen: Policies, Leitlinien, Empfehlungen
- ▶ Medizin-Bibliothek: was tun? Das Beispiel ZB MED
- ▶ Fragen?

Was sind Forschungsdaten?

- ▶ *"Forschungsprimärdaten sind Daten, die im Verlauf von Quellenforschungen, Experimenten, Messungen, Erhebungen oder Umfragen entstanden sind. Sie stellen die Grundlagen für die wissenschaftlichen Publikationen dar. In Abhängigkeit von der jeweiligen Fachzugehörigkeit sind die Forschungsprimärdaten unterschiedlich zu definieren. Die Wissenschaftler sollen in ihren Fachcommunities selber entscheiden, ob bereits Rohdaten hierzu zählen oder ab welchem Grad der Aggregation die Daten langfristig aufzubewahren sind."*

Was sind Forschungsdaten?

- ▶ abhängig von Disziplin
- ▶ Forschungsdaten in der Medizin:
 - Bilddaten aus bildgebenden Verfahren (z.B. MRT),
 - Sensordaten aus Biosignal- oder Vitalparametermessung (z.B. EKG, EEG),
 - Biomaterialdaten aus Laboruntersuchungen (z.B. Blutproben, Genom-Daten),
 - Befunddaten aus ärztlichen Diagnostik (z.B. Anamnese),
 - Statistikdaten (z.B. aus anonymisierten Befunddaten),
 - Klassifikationen und Codes zu Krankheiten oder Materialien (z.B. International Statistical Classification of Diseases and Related Health Problems (ICD)),
 - Stammdaten der Patientenverwaltung (z.B. aus Krankenhausinformationssystemen),
 - ...

Warum „Open Data“?

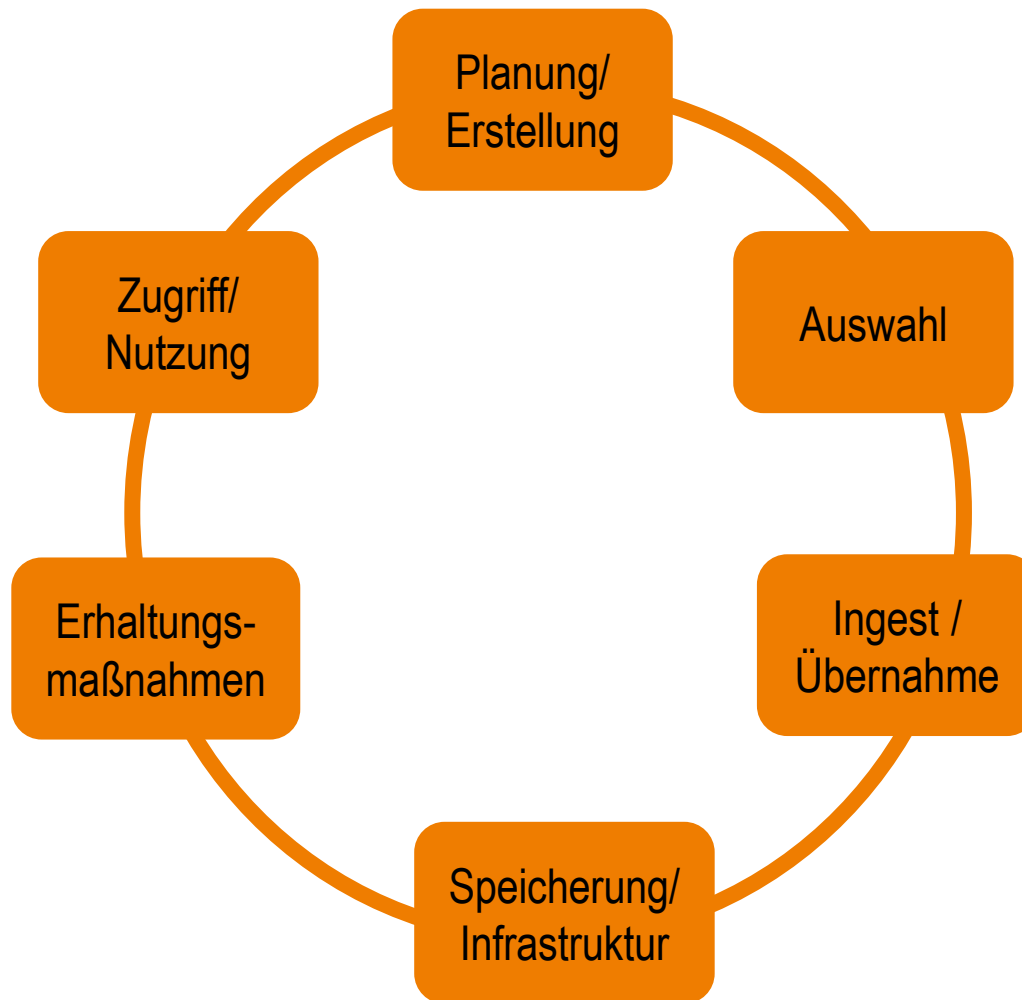
- ▶ Transparenz und Nachprüfbarkeit der Forschungsergebnisse
- ▶ Nachnutzung in neuen Kontexten / weitere Analysen
- ▶ Vermeidung von Doppelarbeit und -finanzierung
- ▶ Sichtbarkeit der Daten → wissenschaftliche Reputation durch Zitationen
- ▶ Erfüllung der Anforderungen von Förderinstitutionen wie DFG und Europäische Kommission oder der eigenen Institution

Definition: Forschungsdatenmanagement

- ▶ *Data Management refers to the storage, access and preservation of data produced from a given investigation. Data management practices **cover the entire lifecycle of the data**, from planning the investigation to conducting it, and from backing up data as it is created and used to long term preservation of data deliverables after the research investigation has concluded.*
- ▶ **Specific activities and issues:**

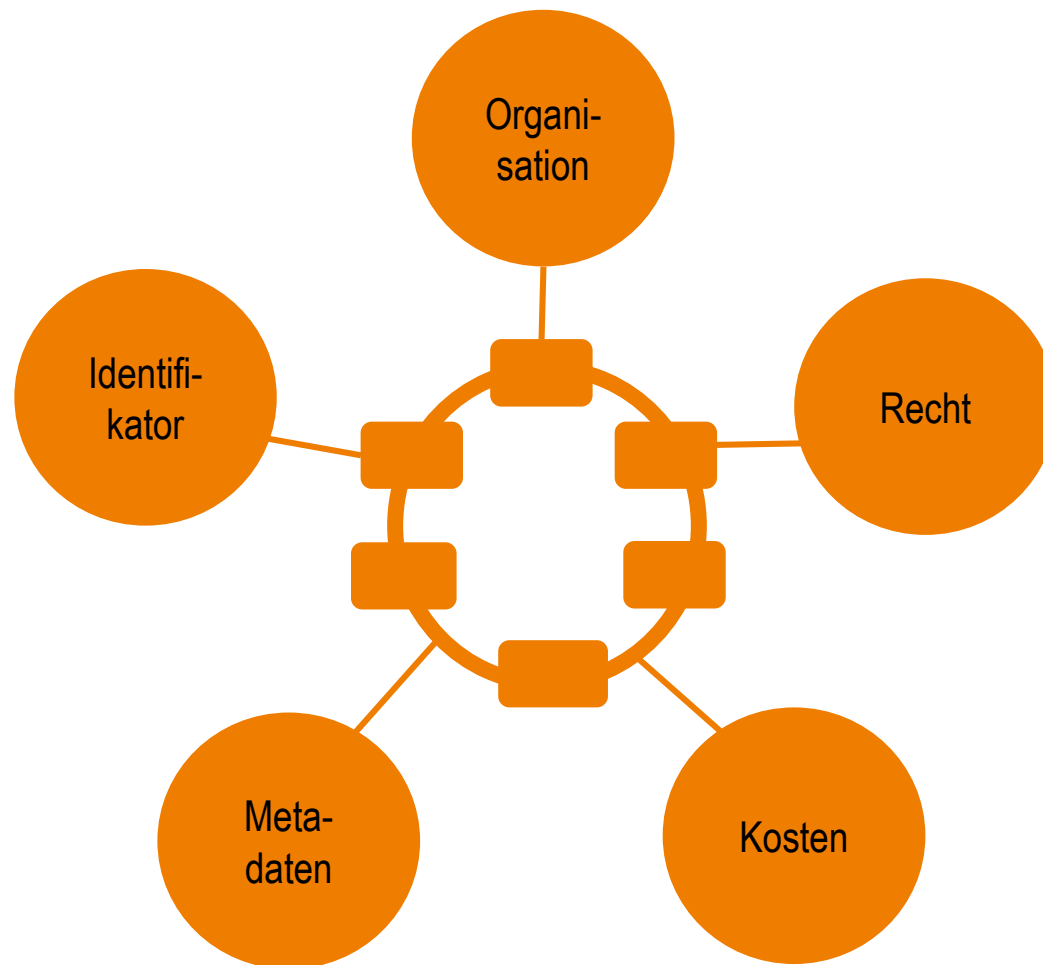
<i>File naming</i>	<i>data security</i>
<i>data quality control and quality assurance</i>	<i>data access</i>
<i>data documentation</i>	<i>data integrity</i>
<i>metadata and controlled vocabularies</i>	<i>data privacy</i>
<i>data storage</i>	<i>data rights</i>
<i>data archiving and preservation</i>	<i>notebook protocols (lab or field)</i>
<i>data sharing and reuse</i>	<i>...</i>

Lebenszyklus von Forschungsdaten



Quelle: Ludwig, J.; Enke, H. (Hrsg.) (2013): Leitfaden zum Forschungsdaten-Management.

Übergreifende Themen bzw. Querschnittsaufgaben im Lebenszyklus von Forschungsdaten



Hemmnisse

- ▶ Vorbehalte der Forschenden
 - Daten können nur für den ursprünglichen Erhebungszweck genutzt werden,
 - viele Daten sind ohne Wert,
 - „nur ich verstehe meine Daten“,
 - Datenschutz verbietet jede Weitergabe,
 - „geht für mein Projekt nicht“,
 - „andere können mehr in den Daten sehen, andere könnten Fehler finden“.
- ▶ Mangelnde Ressourcen für Aufbereitung der Daten
- ▶ Mangelnde Anreize.



Datentypen ... es muss nicht gleich Open Data sein

- ▶ Einzeldaten / Microdata
 - De-identified (ohne IDAT) / absolut / faktisch anonymisiert (z.B. Stat. Bundesamt)
- ▶ Einzeldaten / Beschränkung über Zugriffsart
 - Beschränkte Anfragen / Aggregation on Demand (z.B. I2b2)
- ▶ Aggregierte Daten
 - Beispiel: Ausgewählte Statistiken, Daten in Papier-Publikation
- ▶ (semantische) Metadaten über Daten
 - Beispiel: Working-Instructions, CRF (Case Report Form) Publikation, Biobanken, MetaDataRepositories
- ▶ Metadaten über Sammlungen
 - Beispiel: Studien-Register clinical-trials.gov, Nationale Biobankenregister, biosharing.org, re3data.org

Datenzugriff... es muss nicht gleich Open Data sein

- ▶ Vollzugriff / Besitz: „Datenbesitz“
- ▶ Teilzugang nach Anfrage
- ▶ Modell „Transferstelle“ (z.B. Transferstelle für Daten und Biomaterialienmanagement, Universitätsmedizin Greifswald)
- ▶ Zugang zu faktisch anonymisierten Daten (DeStatis, Statistisches Bundesamt: Gesundheit)
- ▶ Zugang on Screen
- ▶ Zugang remote: „Anfrage einschicken“ (z.B. Informationssystem Versorgungsdaten (Datentransparenz) Daten, DaTraV / DIMDI)
- ▶ Zugang on Site: „Auswertung vor Ort“ (z.B. DaTraV Forscherplatz, Stasi-Unterlagen-Behörde)

Hürden (in der Medizin)

- ▶ Schutz persönlicher Interessen (Personenbezug von Phänotypdaten, *omics-Daten, Biomaterial)
- ▶ Pflicht zur Anonymisierung
- ▶ Gesetzliche Aufbewahrungsfristen (minimal und maximal)
- ▶ Rechtlicher Rahmen: Geflecht aus MBO-Ä (ärztliche Schweigepflicht), Bundes-/ Landesdatenschutzgesetz, Gesetz zur wirtschaftlichen Sicherung der Krankenhäuser und zur Regelung der Krankenhauspflegesätze
- ▶ Urheberrecht
- ▶ Ethische Aspekte bzgl. Erhebung und Nutzung
- ▶ Komplexes Regelwerk bzgl. klinischer Studien („Gute klinische Praxis“)
- ▶ Schutz kommerzieller Interessen (Innovationsschutz)
- ▶ proprietäre Formate

Policies, Leitlinien, Empfehlungen

Rat für Informationsinfrastrukturen (RfII)

- ▶ Gründung im November 2014
- ▶ erstes Positionspapier „Leistung aus Vielfalt – Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland“
- ▶ Empfehlungen mit höchster Priorität:
 - „Aufbau einer Nationalen Forschungsdateninfrastruktur – Kompetenzen bündeln und Grundversorgung mit Services für das Forschungsdatenmanagement schaffen“
 - „Datenmanagementpläne als Planungsinstrumente im Forschungsprozess einsetzen“

Rahmensetzung der Forschungsförderer Deutsche Forschungsgemeinschaft DFG

▶ Empfehlung Nr. 7 (von 17), 1997/2013

„Primärdaten als Grundlagen für Veröffentlichungen sollen auf haltbaren und gesicherten Trägern in der Institution, wo sie entstanden sind, für zehn Jahre aufbewahrt werden.“



Rahmensetzung der Forschungsförderer Deutsche Forschungsgemeinschaft DFG

Seite 1 von 2

▶ Oktober 2015

Deutsche
Forschungsgemeinschaft

Leitlinien zum Umgang mit Forschungsdaten

Forschungsdaten sind eine wesentliche Grundlage für das wissenschaftliche Arbeiten. Die Vielfalt solcher Daten entspricht der Vielfalt unterschiedlicher wissenschaftlicher Disziplinen, Erkenntnisinteressen und Forschungsverfahren. Zu Forschungsdaten zählen u.a. Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Surveydaten, Objekte aus Sammlungen oder Proben, die in der wissenschaftlichen Arbeit entstehen, entwickelt oder ausgewertet werden. Methodische Testverfahren, wie Fragebögen, Software und Simulationen können ebenfalls zentrale Ergebnisse wissenschaftlicher Forschung darstellen und sollten daher ebenfalls unter den Begriff Forschungsdaten gefasst werden. Die langfristige Sicherung und Bereitstellung der Forschungsdaten leistet einen Beitrag zur Nachvollziehbarkeit und Qualität der wissenschaftlichen Arbeit und eröffnet wichtige Anschlussmöglichkeiten für die weitere Forschung. Die Allianz der Wissenschaftsorganisationen hat sich bereits mit den im Jahr 2010 verabschiedeten „Grundsätze[n] zum Umgang mit Forschungsdaten“ für die langfristige Sicherung von, den grundsätzlich offenen Zugang zu und die Berücksichtigung fachdisziplinärer Regularien im Umgang mit Forschungsdaten ausgesprochen.¹ Die „Leitlinien zum Umgang mit Forschungsdaten“ konkretisieren den mit den „Grundsätzen“ vorgegebenen Rahmen im Kontext der DFG-Förderregularien.

Rahmensetzung der Forschungsförderer

Europäische Kommission

- ▶ Die Europäische Kommission setzte mit Horizon 2020 ein Pilotprojekt zur Open-Access-Veröffentlichung von Forschungsdaten auf („Pilot on Open Research Data“) und fordert ausdrücklich dazu auf, Forschungsdaten zu publizieren.
- ▶ Ausweitung des Open Data Research Pilot auf alle Horizon 2020-Programmbereiche im Arbeitsprogramm 2016-2017:

*„Die Projekte, die am Piloten teilnehmen, **müssen** die den Publikationen zu Grunde liegende Daten inklusive Metadaten in einem **Datenrepositorium Open Access** zur Verfügung stellen. Im **Datenmanagementplan (DMP)** des Projekts kann dies flexibel festgelegt werden. Mit der Veröffentlichung wird Dritten das Recht eingeräumt, die Daten zu verwenden und zu verbreiten.“*

Institutionelle Rahmensetzung Forschungsdatenpolicies von Hochschulen

Bielefeld (2011/2013)

*„Grundsätze zu Forschungsdaten an der Universität Bielefeld“ &
„Resolution zum Forschungsdatenmanagement“*

Göttingen (2014)

„Forschungsdaten-Leitlinie der Universität Göttingen (einschl. UMG)“

HU Berlin (2014)

„Grundsätze zum Umgang mit Forschungsdaten an der Humboldt-Universität zu Berlin“

Heidelberg (2014)

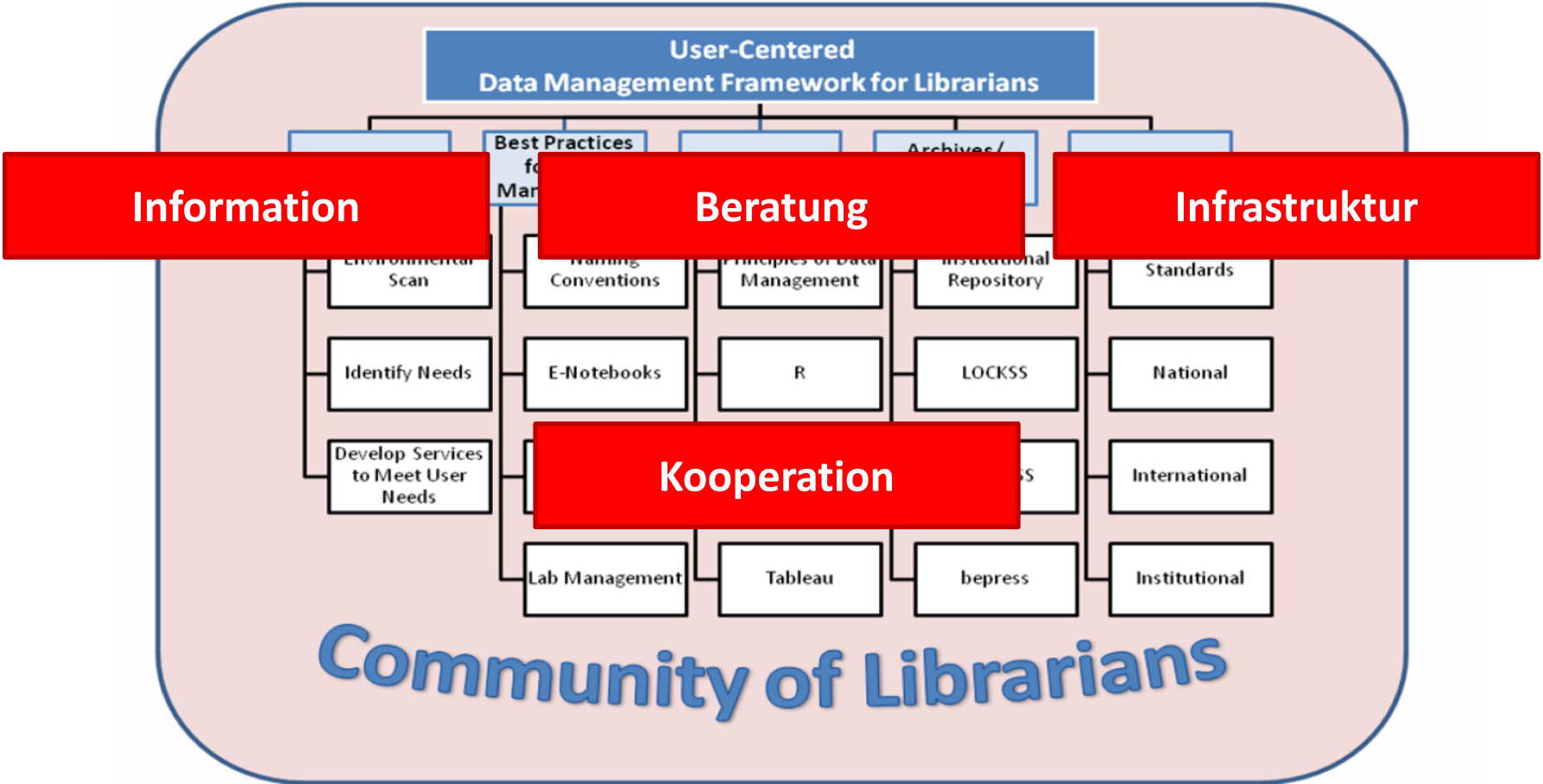
„Research Data Policy“

Medizin / European Medicine Agency EMA

- ▶ Seit 1.1.2015 gilt die

„European Medicines Agency policy on publication of clinical data for medicinal products for human use“
- ▶ First phase: publication of clinical report (“viewing”)
- ▶ Second phase: reviewing various aspects in relation to individual patient data („working with data“)

Forschungsdatenmanagement: Rolle und Aufgaben für Medizinbibliotheken



Quelle: Martin, E. R. (2015): The Role of Librarians in Data Science: A Call to Action. Journal of eScience Librarianship 4(2): e1092., 2015: <http://dx.doi.org/10.7191/jeslib.2015.1092>.

Forschungsdatenmanagement: Rolle und Aufgaben für Medizinbibliotheken

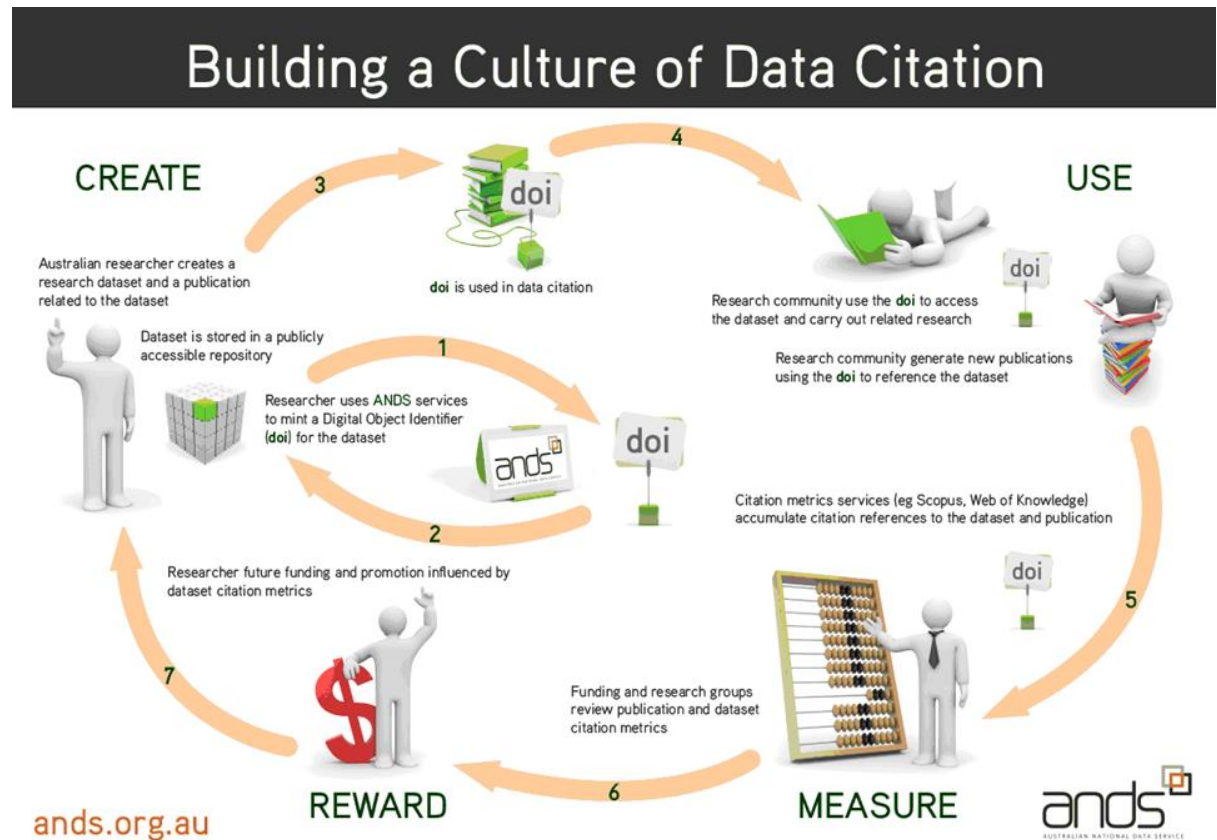
▶ Embedded Data Librarian

„Das Konzept der Embedded Librarianship denkt die Rolle von Bibliothekaren neu. Embedded Librarians begeben sich aus der Bibliothek heraus und hinein in die fachwissenschaftliche Zielgruppe. Dort sind sie – „eingebettet“ – als Mitglied eines Teams direkt und kontinuierlich in alle Abläufe integriert, erlangen so aus erster Hand Wissen über dessen Arbeitsweise, Anforderungen und Bedürfnisse...“

„lokale Forschungsdaten-Support-Teams“

Services ZB MED im Forschungsdatenmanagement: DOI-Service

- ▶ ZB MED ist Mitglied von DataCite und vergibt DOI für Datenzentren aus den Lebenswissenschaften



Quelle: Australian National Data Service: http://www.ands.org.au/_data/assets/pdf_file/0003/383025/data_citation_poster.pdf (zuletzt abgerufen am 23.09.2016).

Lebenswissenschaftliche Forschungsdatenrepositorien

► Tabelle Forschungsdatenrepositorien

[Startseite](#) > [Publizieren](#) > [Forschungsdaten](#) > [Forschungsdatenrepositorien](#)



Forschungsdatenrepositorien

In den folgenden Forschungsdatenrepositorien können Sie – unter Beachtung der fachspezifischen und organisatorischen Rahmenbedingungen der Betreiber – lebenswissenschaftliche Forschungsdaten publizieren. Um eine Auswahl für Sie geeigneter Repositorien zu treffen, ist es möglich, die Tabelle nach den in den Spaltenüberschriften angegebenen Kriterien zu filtern. Drücken Sie dafür auf den nach unten zeigenden Pfeil und wählen Sie die für Sie relevanten Filterkriterien aus.

Aktueller Stand: 08.06.2016

	Alle auswählen ▼	Alle auswählen ▼	Alle auswählen ▼	Alle auswählen ▼	Alle auswählen ▼
Name	Fachlicher lebenswiss. Schwerpunkt	Weiteres Fachgebiet	Regional begrenzte Inhalte	Sitz des Betreibers	Publikation nur in Zusammenhang mit Artikel möglich?
↗ 1000 Functional Connectomes Project	Neurowissenschaften	Medizin	nicht begrenzt	USA	nein
↗ AceView	Biologie	kein	nicht begrenzt	USA	nein
↗ Addgene	Biologie	kein	nicht begrenzt	USA	nein
↗ AEKOS Data Portal (Australian Ecological and Knowledge Observation System Data Portal)	Umwelt	kein	Australien	Australien	nein

<http://www.publisso.de/open-access-publizieren/forschungsdaten/forschungsdatenrepositorien/>

Datenpublikation in German Medical Science

Data from: The active participation of German-speaking countries in conferences of the Association for Medical Education in Europe (AMEE) between 2005 and 2013: a reflection of the development of medical education research?



Files in this package

Content in the Dryad Digital Repository is offered "as is." By downloading files, you agree to the [Dryad Terms of Service](#). To the extent possible under law, the authors have waived all copyright and related or neighboring rights to this data.

Title	German-speaking AMEE participation 2005-2013
Downloaded	10 times
Download	AMEE 2005-2013 Beiträge mit dt. Beteiligung-neu.sav (755.5Kb)
Details	View File Details

When using this data, please cite the original publication:

Raes P, Bauer D, Schöppe F, Fischer MR (2014) The active participation of German-speaking countries in conferences of the Association for Medical Education in Europe (AMEE) between 2005 and 2013: a reflection of the development of medical education research? GMS Zeitschrift für Medizinische Ausbildung 31(3): 28. <http://dx.doi.org/10.3205/zma000920>

Additionally, please cite the Dryad data package:

Raes P, Bauer D, Schöppe F, Fischer MR (2014) Data from: The active participation of German-speaking countries in conferences of the Association for Medical Education in Europe (AMEE) between 2005 and 2013: a reflection of the development of medical education research? Dryad Digital Repository. <http://dx.doi.org/10.5061/dryad.9b56t>

[Cite](#) | [Share](#)

- ▶ ZB MED kooperiert mit DRYAD im Rahmen seines Publikationsportals german medical science GMS.
- ▶ Kostenübernahme durch ZB MED.



Ziele

- ▶ dauerhafte Archivierung von Forschungsdaten aus den Lebenswissenschaften
- ▶ Sicherstellung der Nachnutzbarkeit der Forschungsdaten
- ▶ Publikation singulärer Forschungsdaten („raw research data“) sowie von Forschungsdaten, die mit einem Volltext verknüpft sind („enhanced publication“)
- ▶ Lizenzierung der Daten im Sinne von Open Data, also der Möglichkeit zur Nachnutzung

Forschungsdaten

Titel

Hilfe

Titel* ?

Urheberschaft

Autor* ?

+ - + -

Dateiupload

Medium* ?

CopyrightJahr* ?

Lizenz* ?

Embargo ?

Erschließung

Beschreibung* ?

- ▶ Informationstexte in Form von Frequently Asked Questions (FAQs)
- ▶ Workshops und Vorträge für Wissenschaftlerinnen und Wissenschaftler sowie für Multiplikatoren in wissenschaftlichen Bibliotheken, an Graduiertenschulen und an lebenswissenschaftlichen Fachbereichen von Hochschulen
- ▶ Persönliche Beratung über ein Kontaktformular oder telefonisch; dabei geht es v.a. um Fragen wie: Veröffentlichung von Forschungsdaten und damit zusammenhängende Themen
Erstellung von Datenmanagementplänen, rechtliche Aspekte etc.,
- ▶ Video-Tutorials zu verschiedenen Services im Forschungsdatenmanagement

Zum Schluss...

▶ In memory of data lost..... (Kristin Briney 2015)

▶ hopefully not (ZB MED 2016)

Danke! Fragen?

Birte Lindstädt

Dr. Jasmin Schmitz

Open-Access-Publizieren und
-Beraten

Gleueler Straße 60

50931 Köln

lindstaedt@zbmed.de

schmitz@zbmed.de

