

# FAIRes Datenmanagement mit dem DARIAH-DE Repository

Melina Jander & Lukas Weimer

DHd 2022: Kulturen des digitalen Gedächtnisses

08. März 2022



Digitale Forschungsinfrastruktur für die  
Geistes- und Kulturwissenschaften



# Programm

## 9:00 -10:00 Uhr: Willkommen und Einleitung

- Kurzüberblick DFA
- Vorstellungsrunde
- FAIRes Datenmanagement mit der DARIAH-DE DFA

## 10:00 - 10:15 Uhr: Pause

## 10:15 - 11:30: Hands-On #1

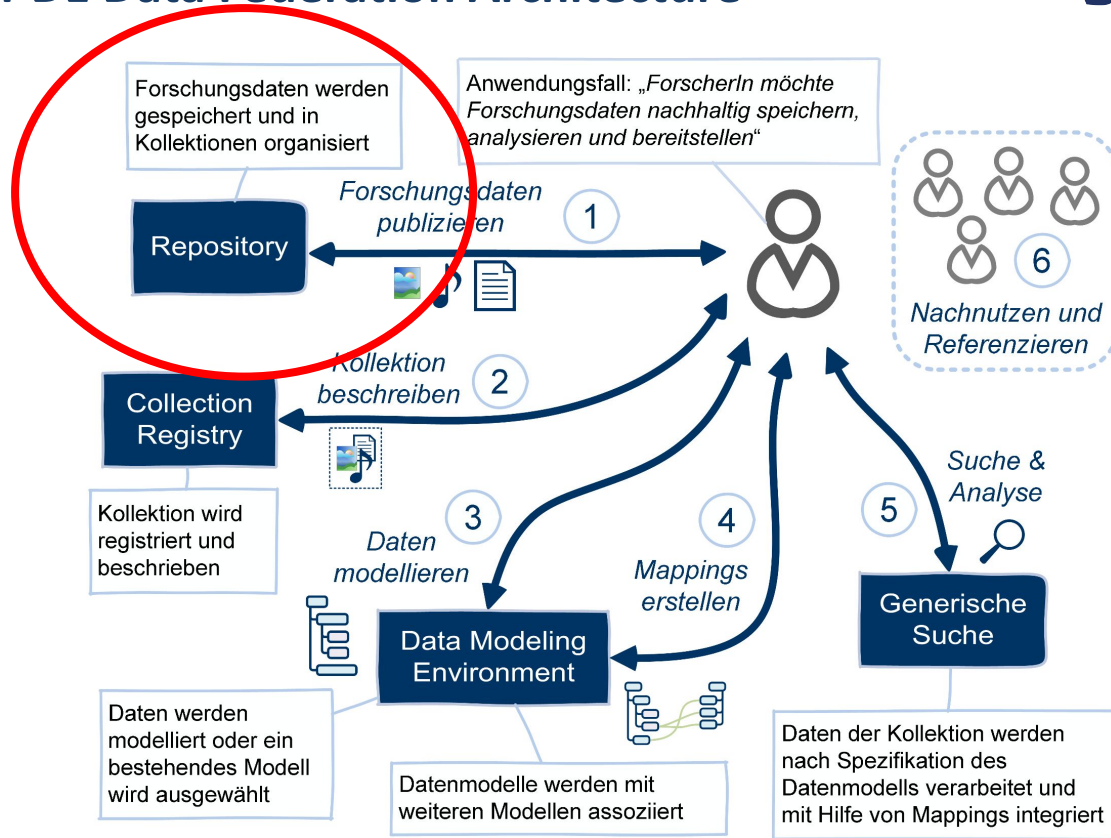
- DARIAH-DE Publikator

## 11:30 - 11:45 Pause

## 11:45 - 13:00: Hands-On #2

- Weitere Elemente der DFA
- Abschluss

# Die DARIAH-DE Data Federation Architecture



# Was bietet DARIAH-DE Nutzenden?

Repository   Publikator



Generische Suche



Beratung

Collection Registry



Forschungsdaten



Veranstaltungen



Geo-Browser



TaDiRAH



Topics Explorer

# Wer sind wir?



Digitale Forschungsinfrastruktur für die  
Geistes- und Kulturwissenschaften



**Melina Jander**

Niedersächsische Staats- und  
Universitätsbibliothek Göttingen

[jander@sub.uni-goettingen.de](mailto:jander@sub.uni-goettingen.de)



**Lukas Weimer**

Niedersächsische Staats- und  
Universitätsbibliothek Göttingen

[weimer@sub.uni-goettingen.de](mailto:weimer@sub.uni-goettingen.de)

Bilder: <https://pixabay.com/de/vectors/m%ac3%a44mlich-weiblich-symbol-1354358/>, 'Roverhale'

## Und wer sind Sie?

...

Wo kommen Sie her?

Was erwarten Sie sich vom Workshop?

# FAIRes Datenmanagement mit der DARIAH-DE DFA

1. Was sind Forschungsdaten?
2. Lebenszyklus von Forschungsdaten
3. Forschungsdaten und gute wissenschaftliche Praxis: Leitlinie 13
4. Was sind die FAIR-Prinzipien?
5. Wieso die DARIAH-DE DFA für FAIRes Forschungsdatenmanagement?

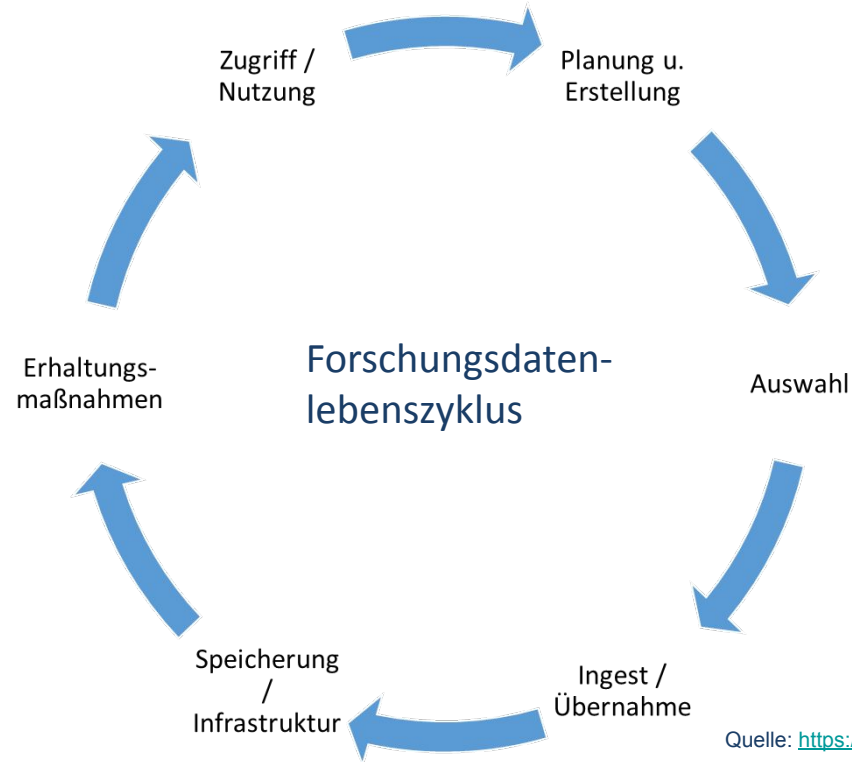
# 1. Was sind Forschungsdaten?

“Forschungsdaten sind nicht allein die (End-)Ergebnisse von Forschung. Es handelt sich vielmehr um **jegliche Daten, die im Zuge wissenschaftlichen Arbeitens entstehen**, z.B. durch Beobachtungen, Experimente, Simulationsrechnungen, Erhebungen, Befragungen, Quellenforschungen, Aufzeichnungen, Digitalisierung, Auswertungen. Zu Forschungsdaten werden auch solche, **nicht selbst gewonnenen Daten, auf die die Wissenschaft zu Forschungszwecken zugreift**, um sie für den konkreten Forschungsprozess als methodisch erforderliche Grundlage zu nutzen.”

Quelle: Rat für Informationsinfrastrukturen: Herausforderung Datenqualität – Empfehlungen zur Zukunftsfähigkeit von Forschung im digitalen Wandel, 2. Aufl. Göttingen 2019, S. B-3. URL: <https://rfii.de/?p=4043>.



## 2. Lebenszyklus von Forschungsdaten



Quelle: <https://de.dariah.eu/weiterführende-informationen>

# 3. Forschungsdaten und gute wissenschaftliche Praxis: Leitlinie 13

## Herstellung von öffentlichem Zugang zu Forschungsergebnissen

“Grundsätzlich bringen Wissenschaftlerinnen und Wissenschaftler **alle Ergebnisse** in den wissenschaftlichen Diskurs ein. **Im Einzelfall** kann es aber Gründe geben, Ergebnisse nicht öffentlich zugänglich (im engeren Sinne in Form von Publikationen, aber auch im weiteren Sinne über andere Kommunikationswege) zu machen; dabei darf diese Entscheidung nicht von Dritten abhängen. Wissenschaftlerinnen und Wissenschaftler entscheiden **in eigener Verantwortung** – unter Berücksichtigung der Gepflogenheiten des betroffenen Fachgebiets –, ob, wie und wo sie ihre Ergebnisse öffentlich zugänglich machen. Ist eine Entscheidung, Ergebnisse öffentlich zugänglich zu machen, erfolgt, beschreiben Wissenschaftlerinnen und Wissenschaftler diese vollständig und nachvollziehbar. Dazu gehört es auch, soweit dies möglich und zumutbar ist, **die den Ergebnissen zugrunde liegenden Forschungsdaten, Materialien und Informationen, die angewandten Methoden sowie die eingesetzte Software verfügbar zu machen** und Arbeitsabläufe umfänglich darzulegen. Selbst programmierte Software wird unter Angabe des Quellcodes öffentlich zugänglich gemacht. Eigene und fremde Vorarbeiten weisen Wissenschaftlerinnen und Wissenschaftler vollständig und korrekt nach.”

Quelle: Kodex “Leitlinien zur Sicherung guter wissenschaftlicher Praxis”. DFG 2019, S. 18-19.

## 4. Was sind die FAIR-Prinzipien?

Auffindbar  
(**F**indable)



Zugänglich  
(**A**ccessible)



Interoperabel  
(**I**nteroperable)

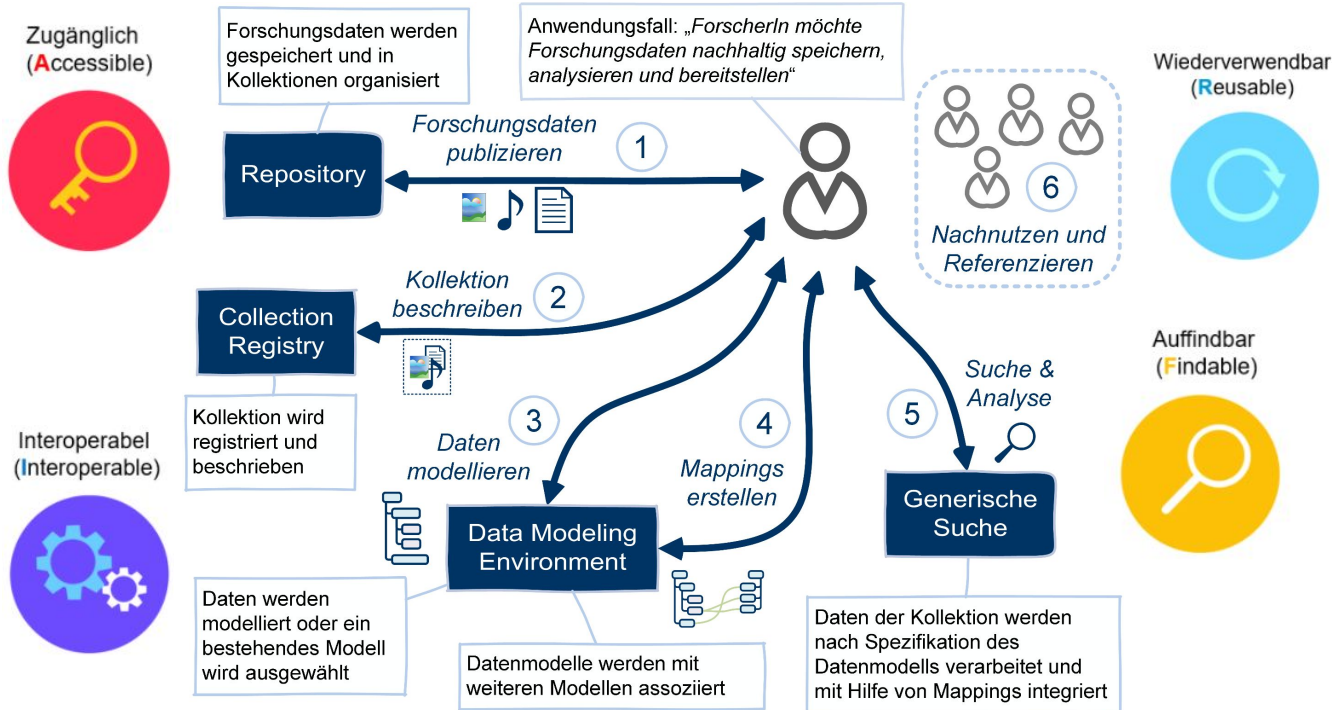


Wiederverwendbar  
(**R**eusable)

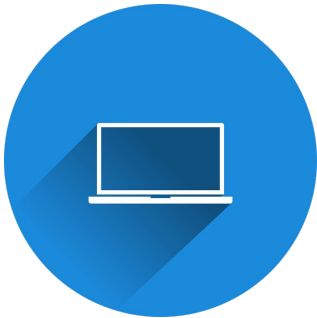


Quelle: Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016).  
<https://doi.org/10.1038/sdata.2016.18>

# 5. Wieso die DARIAH-DE DFA für FAIRes Forschungsdatenmanagement?



# Hands-On #1: Technische Voraussetzungen



## Hands-On #1: DARIAH-DE Publikator

Der Publikator ist ein Werkzeug, das:

- das **Einspielen von Daten** über eine grafische Nutzendenoberfläche (**GUI**) erleichtert,
- deren **Veröffentlichung** vorbereitet,
- die Nutzenden **Schritt für Schritt** durch den Prozess bis zur **persistent referenzierbaren Veröffentlichung** führt:
  - Anlegen einer Datensammlung (Kollektion)
  - Auszeichnen der Kollektion mit Metadaten
  - Anhängen bzw. Hochladen der Dateien u. weiterer Metadaten
  - Publizieren der Kollektion

# Hands-On #1: Beispieldatensatz

## XML-TEI-Bible:

<https://repository.de.dariah.eu/1.0/dhcrud/21.11113/0000-000C-D90A-5>

- 1 übergeordneter Ordner mit
  - 5 untergeordneten Ordnern
  - 1 .txt-Datei
  
- Art der Daten:
  - span. Bibeltexte
  - beschreibende Texte
  - Dateien zur Visualisierung



The screenshot shows the DARIAH-DE repository interface. At the top right, there is a search bar labeled 'REPOSITORY SUCHE' and a 'Sammlung' dropdown menu with 'XML-TEI Bible' selected. The main content area displays 'XML-TEI BIBLE' and 'XML-TEI Bible' with a search icon. Below this, there are 'Links zur Ressource' (Permalink (Handle), Permalink (DOI)) and a 'Datensatz' section. The 'Datensatz' section has a 'DATENSATZ IM QUELLEMODELL' button and a 'HERUNTERLADEN' button. The metadata is displayed in a structured format:

```
► Dc:
► Title: XML-TEI Bible
► Creator: José Calvo Tello
► Subject: Direct speech
► Description: This project contains the biblical text in XML-TEI (66 books). The encoded text is in Spanish, but the codification (elements, attributes, values, ids) is in English. It makes explicit following information: Books, chapters, pericopes and verses. References to peoples, places, times, groups and books, using ids. Direct speech, including who is communicating, to whom and how (written, oral, prayer...).
► Date: 2020-05-28T05:55:18Z
► Type: text
► Format: text/xml
► Identifier: doi:10.20375/0000-000C-D900-2
► Identifier: hdl:21.11113/0000-000C-D900-2
► Language: es
► Relation: hdl:21.11113/0000-000C-D90A-5
► Rights: CC-BY
```

## Hands-On #2: Weitere Module der DFA

Publikator u. Repository	<ul style="list-style-type: none"> <li>• Einspielen von Forschungsdaten über GUI</li> <li>• Auszeichnen mit Metadaten</li> <li>• Publizieren der Daten u. Metadaten</li> </ul>
Repository Suche	<ul style="list-style-type: none"> <li>• Suchen u. Auffinden von im Repository abgelegten Daten</li> </ul>
Collection Registry	<ul style="list-style-type: none"> <li>• Registrierung von Kollektionen</li> <li>• Auszeichnen mit Metadaten</li> </ul>
Data Modeling Environment	<ul style="list-style-type: none"> <li>• Modellierung u. Assoziation von Daten</li> </ul>
Generische Suche	<ul style="list-style-type: none"> <li>• Front-End für die in der Collection Registry u. dem Repository verfügbar gemachten Daten u. deren Metadaten</li> </ul>





## Hands-On #2: Collection Registry

- Registry für sowohl veröffentlichte und digital zugängliche Sammlungen (Kollektionen) als auch nicht-öffentliche und nur analog verfügbare Sammlungen
- Auszeichnen der Sammlungen mit Metadaten
- Schema: DARIAH Collection Description Data Model (DCDDM)



The screenshot shows the DARIAH-DE Collection Registry interface. At the top left is the DARIAH-DE logo and name. To its right is the text 'COLLECTION REGISTRY'. Below the logo, the text 'Collection Registry' is displayed. Underneath, there are two bullet points: '227 Sammlungsbeschreibungen' and '266 Akteure'. The main content area features a large, dense cluster of small circular icons, each representing a collection or actor. On the right side, there is a section titled 'Letzte Aktivität' (Latest Activity) which lists two recent events: 'Klassifikation von Tragödien und Komödien bei Calderón de la Barca' and 'Zotero-Workshop von CLARIAH-DE, 27./28.05.2021'. Each event entry includes a date and time, a unique identifier, an email address, and a status indicator (e.g., 'Sammlung gespeichert' or 'Sammlung genehmigt').

## Hands-On #2: Generische Suche

- Durchsucht die Collection Registry
- Zwei Such-Modi:
  - Einfache Suche
  - Erweiterte Suche (für detaillierte Suchanfragen)

61 Kollektionen | 1.572.116 Dokumente



Erweiterte Suche



# Zusammenfassung und Nachbesprechung

**Zusammenfassung**

**Diskussion**

**Nachbesprechung**



SUB

NIEDERSÄCHSISCHE STAATS- UND  
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

## Vielen Dank für Ihre Teilnahme!

Melina Jander & Lukas Weimer

DHd 2022: Kulturen des digitalen Gedächtnisses

08. März 2022

 **DARIAH-DE**

Digitale Forschungsinfrastruktur für die  
Geistes- und Kulturwissenschaften



GEORG-AUGUST-UNIVERSITÄT  
GÖTTINGEN