



# Forschungsorientierte Modellierung und Korrelation von Daten in der Föderationsarchitektur von DARIAH-DE

Tobias Gradl

Otto-Friedrich-Universität Bamberg, Medieninformatik

Stefan Schmunk

Staats- und Universitätsbibliothek Göttingen

Bern, 13. Februar 2017

 @DARIAHde

- Forschungsdaten in den Geistes- und Kulturwissenschaften
- Forschungsdaten Föderationsarchitektur
- Schlussbemerkungen

## Definition Forschungsdaten

*"Unter digitalen geistes- und kulturwissenschaftlichen Forschungsdaten werden all jene **Quellen/Materialien und Ergebnisse** verstanden, die im Kontext einer geistes- und kulturwissenschaftlichen Forschungsfrage **gesammelt, erzeugt, beschrieben und/oder ausgewertet** werden und in **maschinenlesbarer Form** zum Zwecke der **Archivierung, Zitierbarkeit** und zur weiteren **Verarbeitung** aufbewahrt werden können."*

- Metadaten, bibliographische Daten, Findbücher
- Digitale und / oder digitalisierte Daten und / oder digitale Repräsentationen von analogem Material
- Digitale Objekte
- Volltexte, Transkriptionen
- Annotierte Volltexte
- Bilder, Filme, Musik & Noten
- Authority Files, Kontrollierte Vokabularien, Ontologien
- etc.



## Wissenschaftliche Sammlungen – epidat

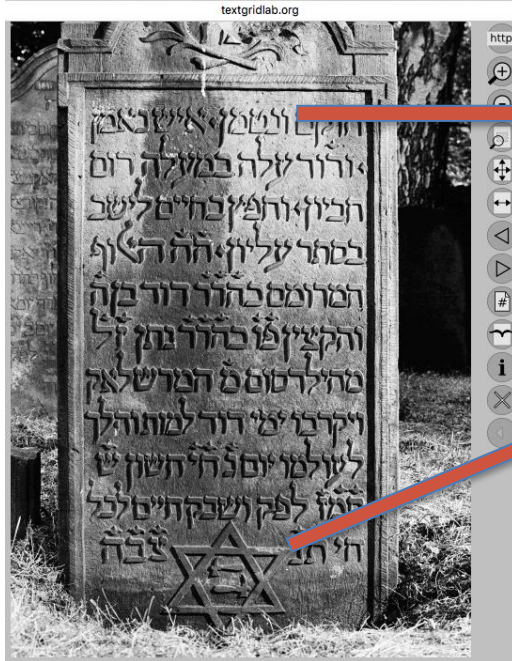
- **Daten:** 30.000 Inschriften von mehr als 130 jüdischen Friedhöfen in Deutschland und den Niederlanden (1050-2000)
- **Enthält:** Digitalisierte Daten, Transkriptionen, Karten, Fotografien und weiteres Archivmaterial
- **Methode:** Visual Analytics, Named-Entity-Recognition, Raum-Zeit-Visualisierungen

### \*File and encoding format:

- data format
  - images: jpg png gif tiff eps
  - maps: svg pdf AutoCad kml
  - text
    - extern: x(html)
    - extern: TEI4 XML (using: epidoc: Epigraphic Documents in TEI XML)
    - extern (wip): TEI5 XML (using: example)
    - extern: plain/text (no format)
    - extern (on demand): PDF
    - intern: TUSTEP file with epidat structure
- metadata format
  - TEI header
  - DC (simple, example)



Source: <http://www.steinheim-institut.de/cgi-bin/epidat?info=edv&art=art6>



Symbole

Fotografien

Grablagen

Weiterführende Informationen zu Epidat:  
 Thomas Kollatz  
[kol@steinheim-institut.org](mailto:kol@steinheim-institut.org)

David (genannt Hammerschlag) ben Natan [26.10.1686] Seitenanfang

Edition und Übersetzung

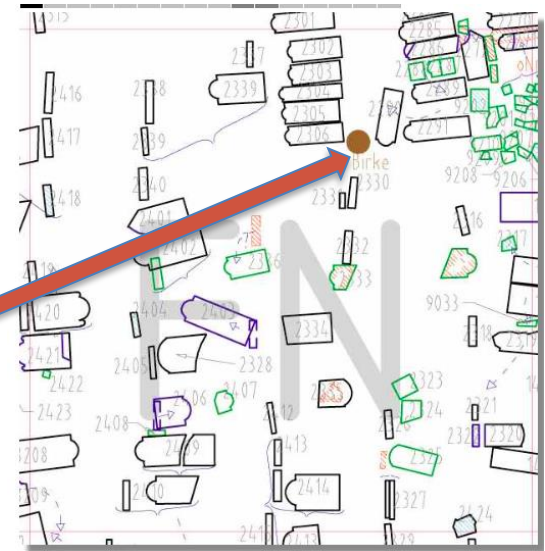
<p>[ס] הוּקַם וּנְטָמָן : אִישׁ נָאֵמָן          : וְדוֹד עָלָה בַּבְּשֵׁלָה רוּם          חֲבוּיוֹן וְחֲפִיץ בְּחַיִּים לִישֵׁב          בְּסִתְרֵי עֲלִיוֹן : הִיָּה הָאֱלֹהִים          הַמְרוֹמֵם כְּהַרְיָד דּוֹד בֶּן הַיְיָ          וְהַקְצִין פִּי כְּהַרְיָד נְתוּן דִּל          מֵהִילְדוּסָם מִי הַמְרִשְׁלָאָק          וְיִקְרָבוּ יָמֵי דּוֹד לְמוֹת הַלֵּךְ          לְעוֹלָמוּ יוֹם גַּי חַיִּי חֲשׁוֹן שִׁי          תִּמְיָו לְפִקֵּי וְשִׁבְקֵי חַיִּים לְכָל          חַי תִּנְצַבְיָהּ</p>	<p><i>Hier errichtet (ein Mal) und ward geborgen ein getreuer Mann, »und David stieg hinauf: hoch ins Versteck, »der, der Leben begehrt, »wird sitzen im Schutz des Höchsten«, es ist der Vornehme, der Erhabene, der geehrte Meister, Herr David, Sohn des Vornehmen und Einflußreichen, Vorstehers und Leiters, des geehrten Meisters, Herrn Natan, sein Andenken zum Segen, aus Hildesheim, genannt Hammerschlag. »Und es nahten Davids Tage dem Tod« und »er ging hin in seine Welt« (am) Tag 3, 18. Cheschan des Jahres 447 der kleinen Zählung, »und er ließ das Leben wie alles Lebende«. Seine Seele sei eingebunden in das Bündel des Lebens</i></p>
---	---

Zl 3: 2Sam 15,30 Zl 4: Ps 34,13 Zl 4f: Ps 91,1 Zl 9: 1Kön 2,1 Zl 9f: Koh 12,5 Zl 11f: nach bBer 61b

## Kritische Edition – Synoptische Ansicht d. Inschriften

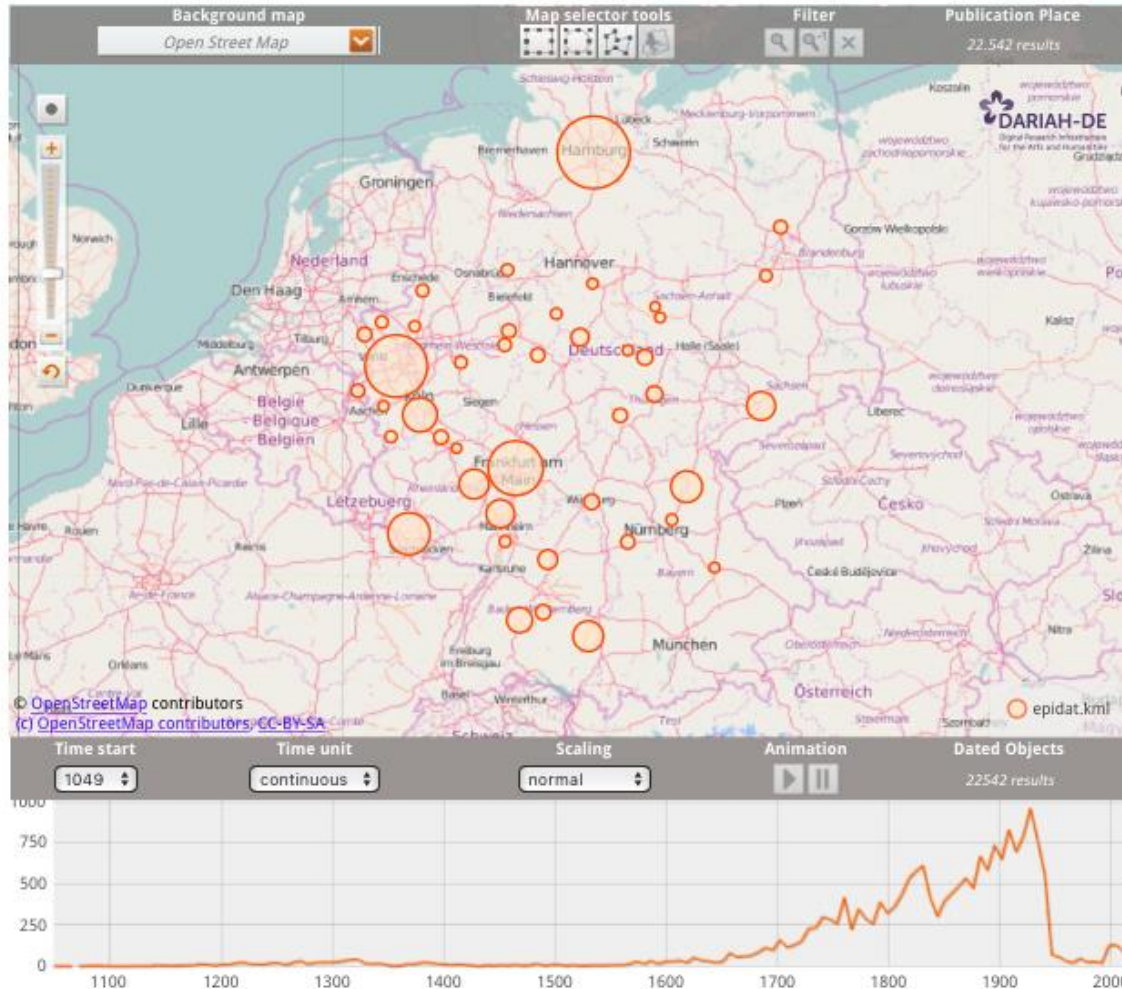
belegt Planquadrat FN

	P	O	N	M	L	K	I	H	G	F	E	D	C	B	A
A		NA	MA	LA											
B		NB	MB	LB	KB	IB									
C		NC	MC	LC	KC	IC	HC	GC							
D		OD	ND	MD	LD	KD	ID	HD	GD	FD	ED	DD			
E		PE	OE	NE	ME	LE	KE	IE	HE	GE	FE	EE	DE		
F		PF	OF	NF	MF	LF	KF	IF	HF	GF	FF	EF	DF	BF	AF
G		PG	OG	NG	MG	LG	KG	IG	HG	GG	FG	EG	DG	BG	AG
H		PH	OH	NH	MH	LH			HH	GH	FH	EH	DH	CH	BH
I									HI	GI	FI	EI	DI	BI	AI
K									HK	GK	FK	EK	DK	CK	BK
L									HL	GL	FL	EL	DL		
M									HM	GM	FM	EM	DM	CM	BM
N									GN	FN	EN	DN	CN	BN	
O									GO	FO	EO				
P									GP	FP	EP				
Q									QP	FP	EQ				
R									GR	FR					





Das DARIAH-DE spatio-temporal interface visualisiert die zeitlich-räumliche Belegung jüdischer Friedhöfe auf Grundlage der datierten epidat Inschriften.



Visualisierung der  
Wissenschaftlichen  
Sammlungen (epidat)

DARIAH-DE Geo-Browser  
[geobrowser.de.dariah.eu](http://geobrowser.de.dariah.eu)

Zeit-Räumliche-  
Kontextualisierung

Quelle: <http://www.steinheim-institut.de/cgi-bin/epidat?info=e4d&lang=de>

## Fragestellung

Wie wurde ein bestimmtes literarisches Werk vom historischen Publikum verstanden?

- Welche Kontexte wurden mit ihm verknüpft?
- Wie wurden Leerstellen gefüllt?
- Mit welchen Gattungskonventionen wurde es in Beziehung gesetzt?
- Nach welchen Wertungskriterien wurde es beurteilt?
- usw.

Quelle: [www.katjamellmann.de](http://www.katjamellmann.de)

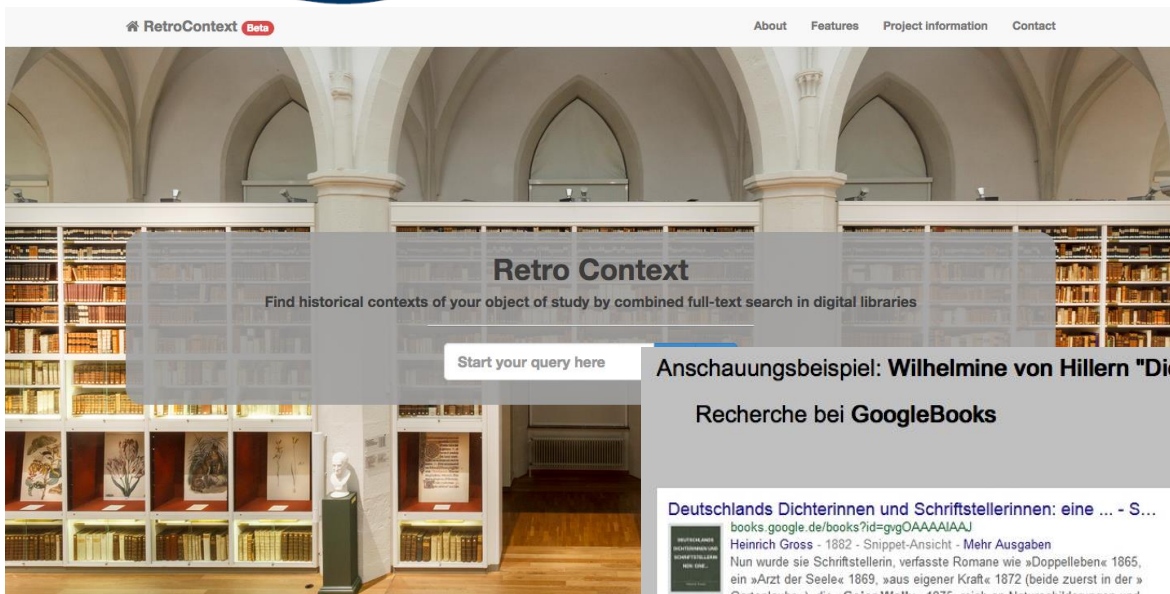
## Quellen

- Rezensionen
- Werkerwähnungen in anderen Kontexten, v.a. Zeitschriftenartikel
- persönliche Rezeptionszeugnisse (Briefe, Tagebücher, Memoiren)
- Bibliothekskataloge (weltweit)

## Recherchehilfen

- Bibliographien
- Namensregister (in Werk- und Briefeditionen u.Ä.)
- **Retrodigitalisate** von Zeitschriften und anderen hist. Drucken





Quelle: [www.retrocontext.de](http://www.retrocontext.de)

## Anschauungsbeispiel: Wilhelmine von Hillern "Die Geier-Wally" (1875)

### Recherche bei GoogleBooks

#### Deutschlands Dichterinnen und Schriftstellerinnen: eine ... - S...

books.google.de/books?id=gvgOAAAAIAAJ  
 Heinrich Gross - 1882 - Snippet-Ansicht - Mehr Ausgaben  
 Nun wurde sie Schriftstellerin, verfasste Romane wie »Doppellebene« 1865, ein »Arzt der Seele« 1869, »aus eigener Kraft« 1872 (beide zuerst in der »Gartenlaube«), die »Geier-Wally« 1875, reich an Naturschilderungen und Seelenmalerei, ...

#### Lexikon der deutschen Nationalliteratur: die deutschen ... - S...

books.google.de/books?id=RivHAAAAQAAJ  
 Adolf Stern - 1882 - Keine Vorschau - Mehr Ausgaben

#### Deutsche Revue - Band 11, Teile 3-4

books.google.de/books?id=ko0mAQAAIAAJ  
 1886 - Snippet-Ansicht - Mehr Ausgaben

#### Deutscher bühnen-almanach - Seite 136

books.google.de/books?id=ZUOAAAAMAAJ  
 1884 - Snippet-Ansicht - Mehr Ausgaben  
 Hr. Schüherl. Neu aufgeführte Stücke: Müllers Vaterfreuden. Familie Schneck. Eine mit Talent. Geier-Wally. Rattenfänger v. Hameln. Auf der Rar. Kaiser Josef im Volke. Böhm in Amerika. Josef Lanner. Lustige Krieg. 10 Mädchen u. kein Mann.

### Im Buch

Geier-Wally

19 Seiten stimmen mit dem Suchbegriff "Geier-Wally" in diesem Buch überein.

Seite 171

... seine gegen die ...  
 derselben sagte er: die **Analyse mit den Mischlingen** sei der Autorin vermutlich erst im Laufe der Arbeit gekommen und sie habe dieselbe glücklich verwertet. Ich vermute, daß Paul Lindau als er diesen glänzenden Esaj geschrieben, sich überhaupt gar nichts dabei gedacht hat, hat er aber vielleicht sich in der Mission genügt, daß die Verfasserin der „Geierwally“ ihre geheimsten Ideen in einer

Seite 324

...  
 genauer erörtert. — Die Neurostik kam während des individuellen Lebens erworben oder durch Vererbung überkommen sein; in letzterem Falle äußert sie sich nicht selten schon in der Kindheit als geistige Frühreife und unstillbarer Wissensdurst. Derartige vererbte, von Geburt an dem Menschen innewohnende Krankheitsanlagen kann der Arzt nicht vollständig beseitigen, aber er und eine gute Ge-

Seite 325

...  
 mung Ihres eminenten Darstellungstalents noch Schöpfungen vollenden werden, die an Originalität und gewaltiger Plastik alle Ihre früheren Leistungen überstrahlen müssen. Ihre Schöpfung muß bei allen, die Verständnis für Poesie und Gestaltungstalent haben, entzückenden Anklang finden, wie sie in gewissen Kreisen andererseits auch vielfach verdammt und verurteilt werden wird. Sie haben mit

Wo ist der Rest dieses Buches?

Ergebnisse 1-3 von 19

# What is the situation for the Digital Humanities?

**Research data** are representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research or scholarship”

C.L. Borgman (2015). Big Data, Little Data, No Data: Scholarship in the Networked World.

Neue Herausforderungen für die Nutzung von Digitalen Forschungsdaten in den Geisteswissenschaften:

- Daten und Sammlungen finden
- Zugriff auf Daten
- Neue und umfangreiche Suchstrategien
- Nutzung von Datenwerkzeugen (Datengenerierung)
- Verlinkung von Artikel mit Forschungsdaten



Werkzeuge werden benötigt für ...

... (Generische) Suchanfragen

... Erstellung und Definition von Metadaten Formaten

... Generierung von (maschinenlesbaren) Crosswalks

... Beschreibung von Forschungsdaten und Sammlungen

... Speicherung und Langzeitarchivierung von Forschungsdaten

Werkzeuge werden benötigt für ...

Wie und wo finde ich  
Forschungsdaten  
für meine Forschung?

# Aufbau der Föderationsarchitektur

*ForscherIn möchte Daten einer  
Sammlung, kontextualisieren,  
analysieren und durchsuchen.*



 **Sammlung  
beschreiben**

*Aufgaben und  
Unterstützungspotenzial*

 **Daten  
definieren**

 **Mappings  
erstellen**

 **Suche**



# Aufbau der Föderationsarchitektur

Die Sammlung wird eingetragen und die Zugriffsschnittstelle auf Daten der Sammlung beschrieben.

ForscherIn möchte Daten einer Sammlung, kontextualisieren, analysieren und durchsuchen.

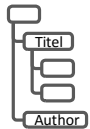
Collection Registry



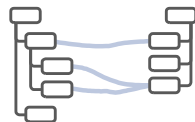
**Sammlung beschreiben**



*Aufgaben und Unterstützungspotenzial*



**Daten definieren**



**Mappings erstellen**



**Suche**

# Aufbau der Föderationsarchitektur

*ForscherIn möchte Daten einer Sammlung, kontextualisieren, analysieren und durchsuchen.*

*Die Sammlung wird eingetragen und die Zugriffsschnittstelle auf Daten der Sammlung beschrieben.*

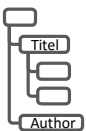
Collection Registry



**Sammlung beschreiben**

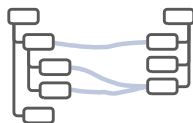


*Das Schema der Daten wird ausgewählt oder neu angelegt.*



**Daten definieren**

Schema Registry

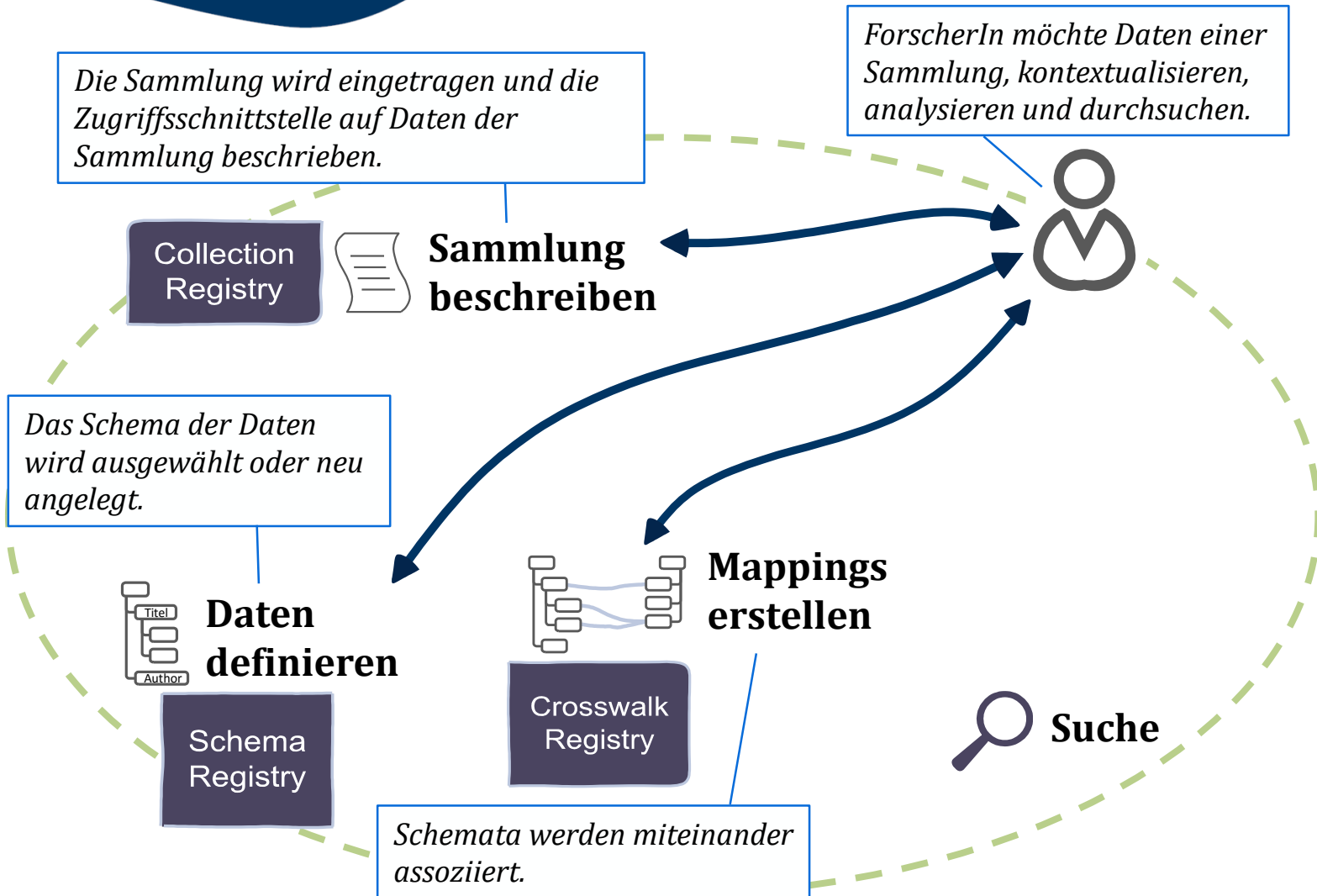


**Mappings erstellen**

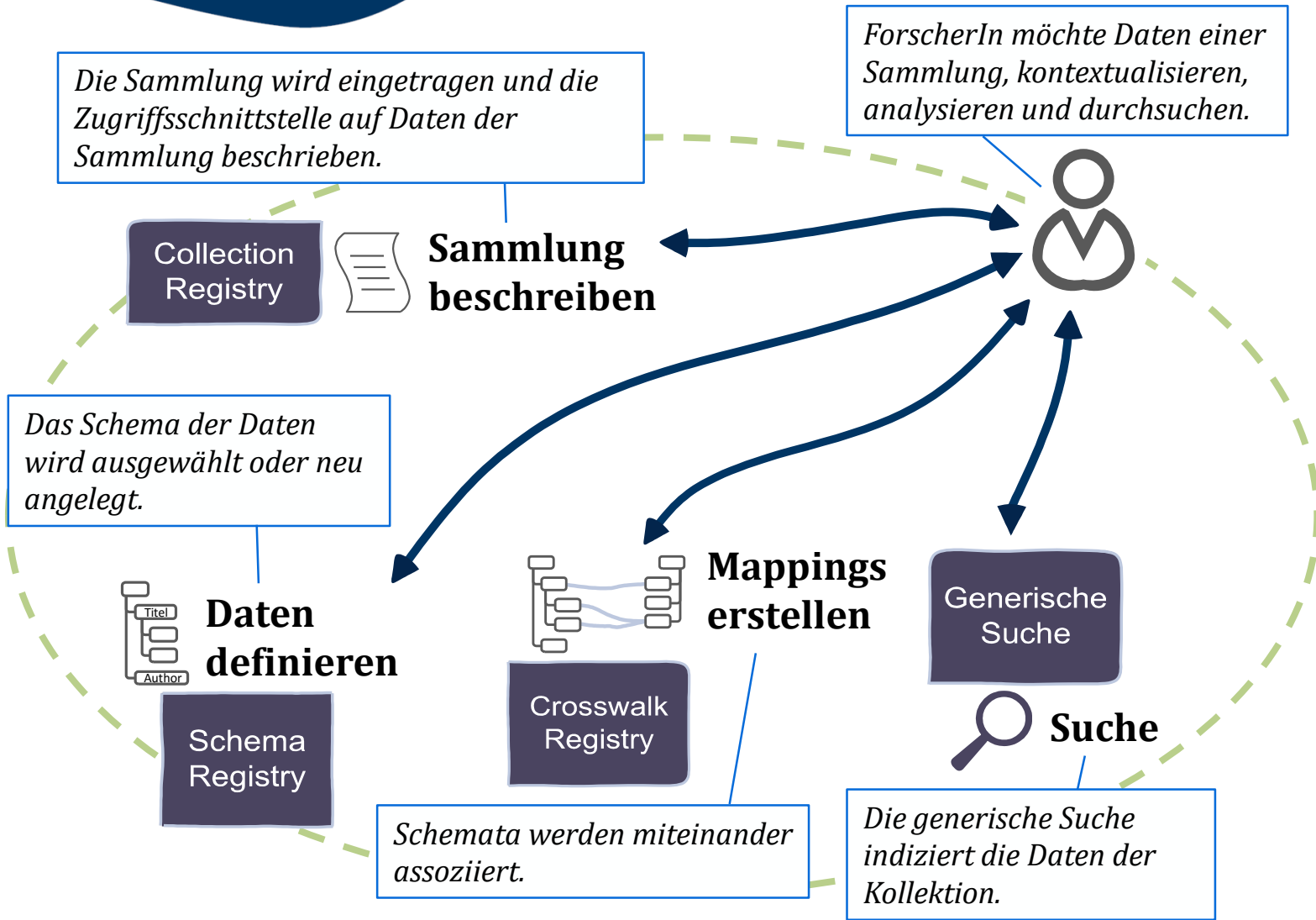


**Suche**

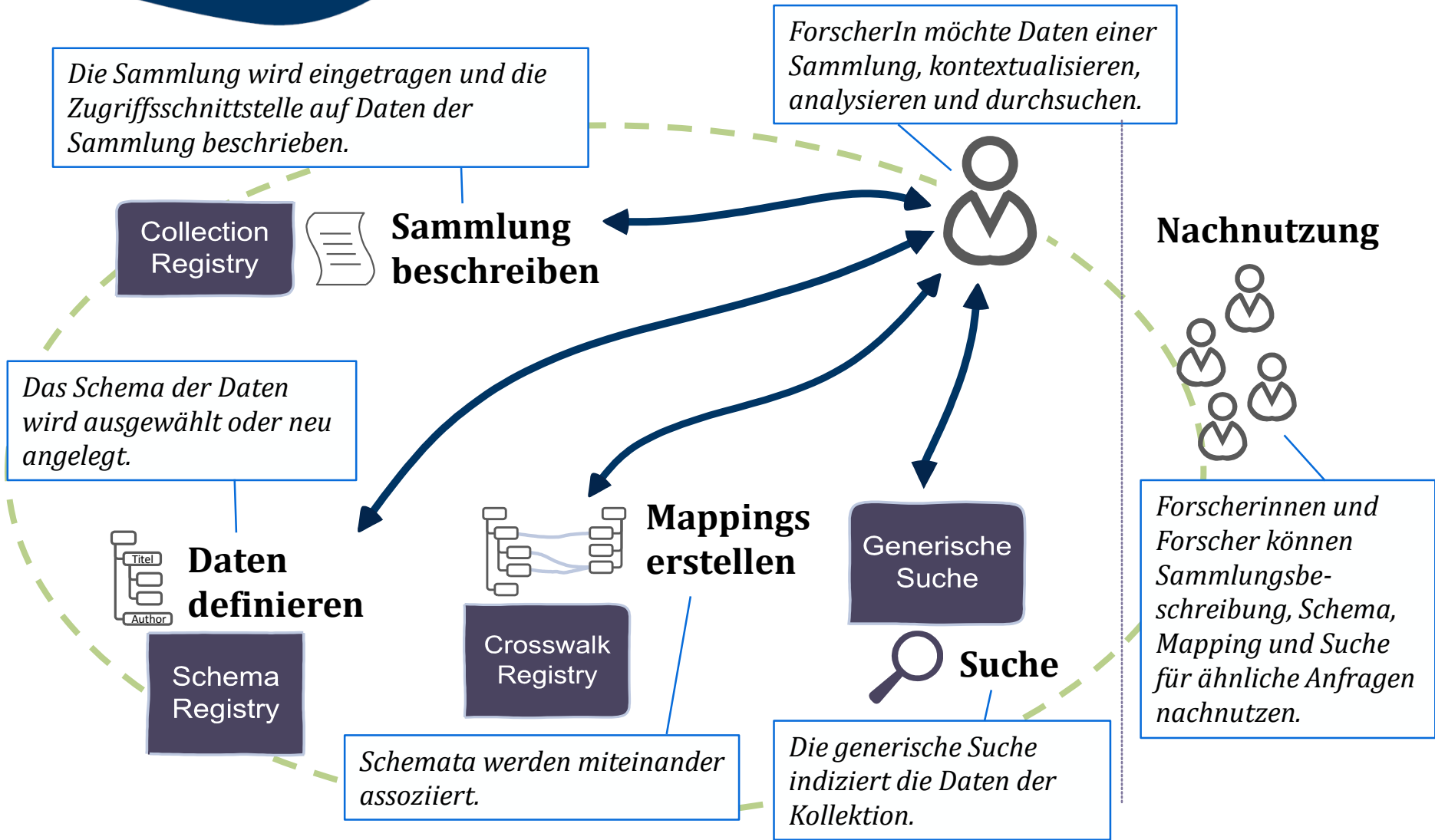
# Aufbau der Föderationsarchitektur



# Aufbau der Föderationsarchitektur

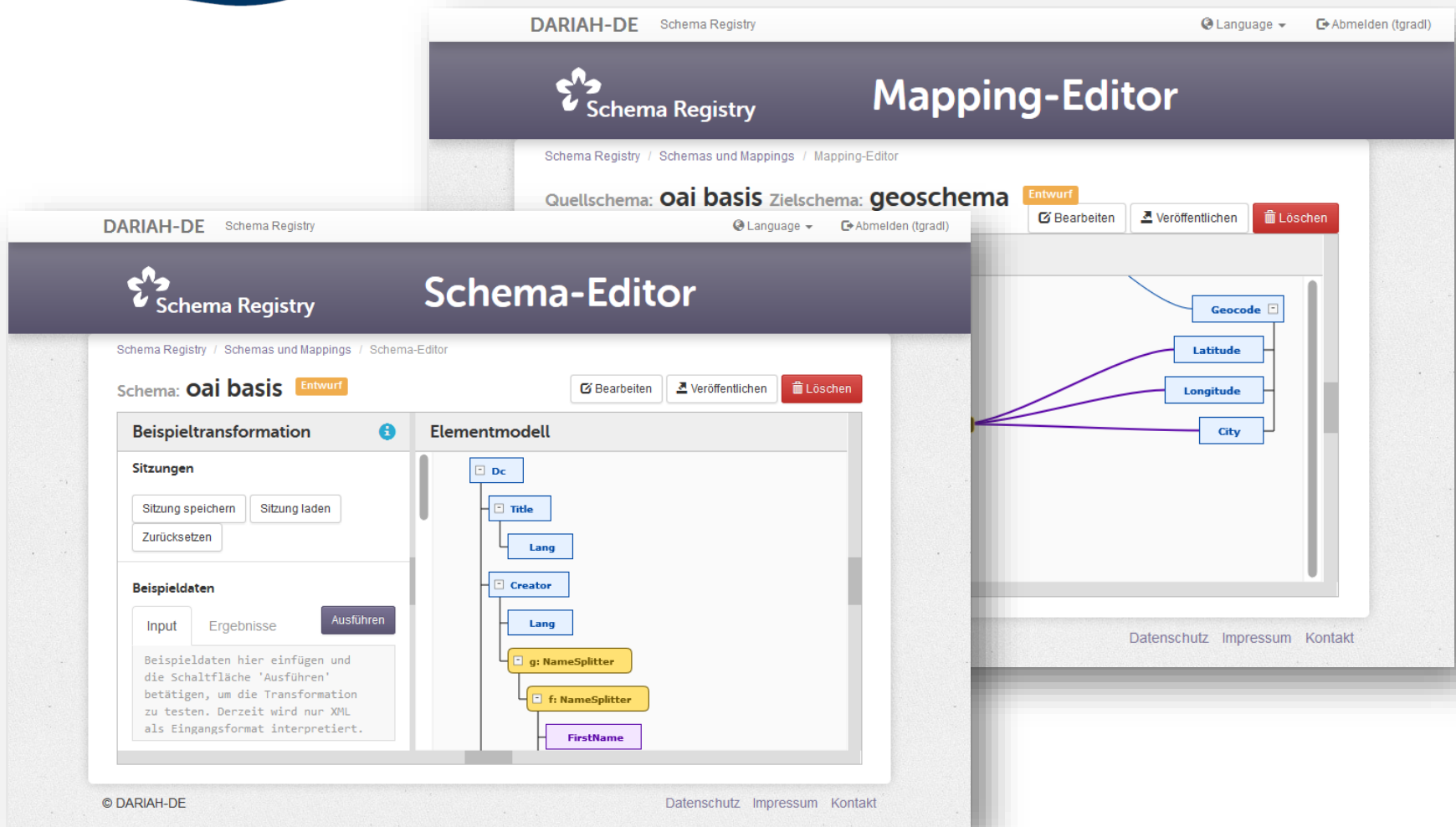


# Aufbau der Föderationsarchitektur









The image displays two overlapping screenshots of the DARIAH-DE Schema Registry web application.

**Top Screenshot: Mapping-Editor**

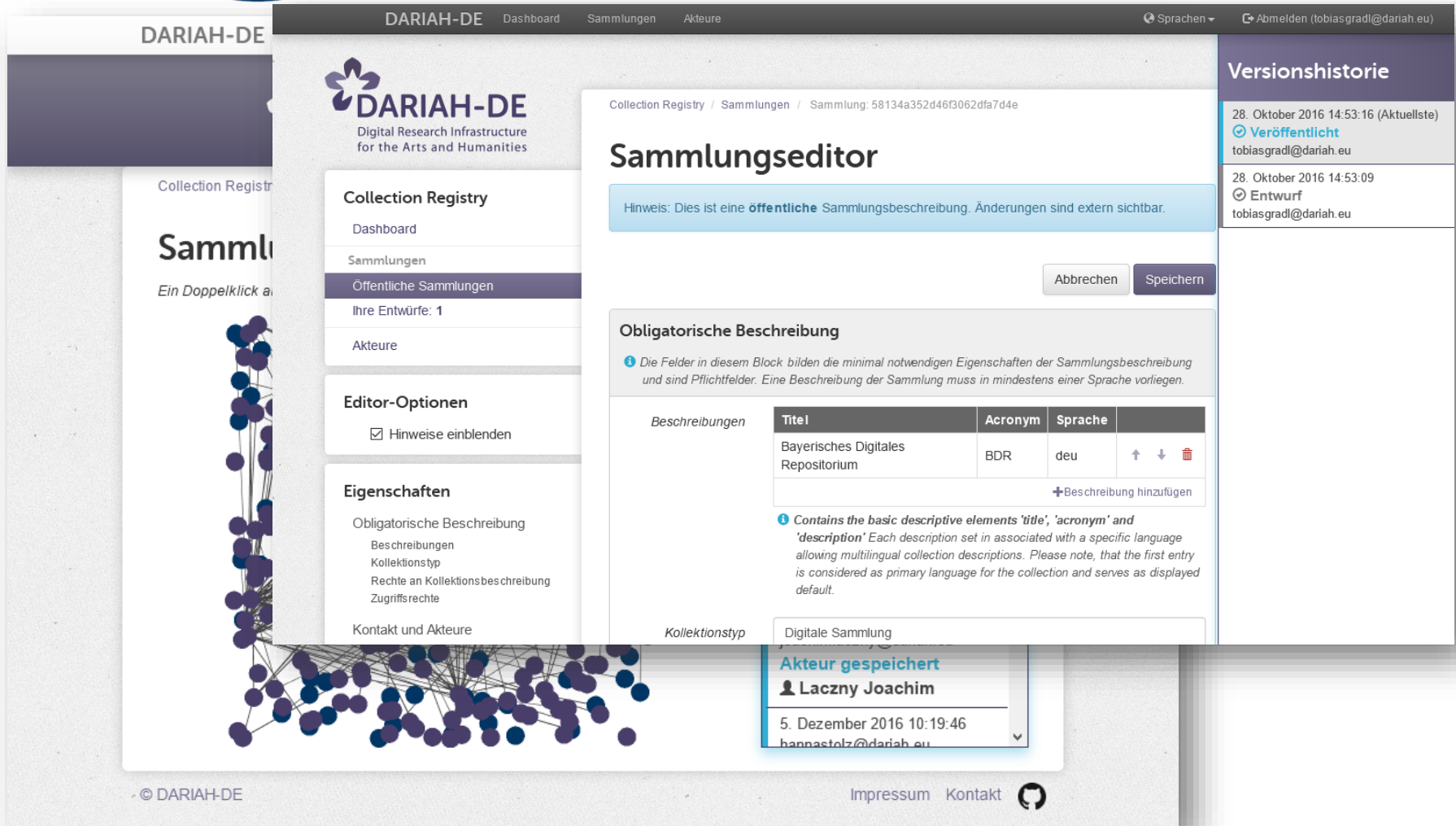
- Header: DARIAH-DE Schema Registry, Language, Abmelden (tgradl)
- Page Title: Mapping-Editor
- Breadcrumbs: Schema Registry / Schemas und Mappings / Mapping-Editor
- Source/Target: Quellschema: **oai basis** Zielschema: **geoschema** (Entwurf)
- Actions: Bearbeiten, Veröffentlichen, Löschen
- Diagram: A mapping diagram showing a source element on the left connected by purple lines to four target elements on the right: Geocode, Latitude, Longitude, and City.
- Footer: Datenschutz Impressum Kontakt

**Bottom Screenshot: Schema-Editor**

- Header: DARIAH-DE Schema Registry
- Page Title: Schema-Editor
- Breadcrumbs: Schema Registry / Schemas und Mappings / Schema-Editor
- Schema: **oai basis** (Entwurf)
- Actions: Bearbeiten, Veröffentlichen, Löschen
- Left Panel:
  - Beispieltransformation**
    - Sitzungen: Sitzung speichern, Sitzung laden, Zurücksetzen
    - Beispieldaten: Input, Ergebnisse, Ausführen
    - Text: Beispieldaten hier einfügen und die Schaltfläche 'Ausführen' betätigen, um die Transformation zu testen. Derzeit wird nur XML als Eingangsformat interpretiert.
- Right Panel: **Elementmodell**
  - Root: Dc
    - Title
    - Lang
    - Creator
      - Lang
      - g: NameSplitter
        - f: NameSplitter
          - FirstName

- Footer: © DARIAH-DE, Datenschutz Impressum Kontakt

Source: <https://schereg.de.dariah.eu>



The screenshot shows the 'Sammlungseditor' (Collection Editor) interface. The main content area is titled 'Sammlungseditor' and contains a warning: 'Hinweis: Dies ist eine öffentliche Sammlungsbeschreibung. Änderungen sind extern sichtbar.' Below this is the 'Obligatorische Beschreibung' section, which includes a table of descriptions and a dropdown for 'Kollektionstyp' (Collection Type).

**Collection Registry**

- Dashboard
- Sammlungen
- Öffentliche Sammlungen
- Ihre Entwürfe: 1
- Akteure

**Editor-Optionen**

- Hinweise einblenden

**Eigenschaften**

- Obligatorische Beschreibung
  - Beschreibungen
  - Kollektionstyp
  - Rechte an Kollektionsbeschreibung
  - Zugriffsrechte
- Kontakt und Akteure

**Sammlungseditor**

Collection Registry / Sammlungen / Sammlung: 58134a352d46f3062dfa7d4e

Hinweis: Dies ist eine **öffentliche** Sammlungsbeschreibung. Änderungen sind extern sichtbar.

Abbrechen Speichern

**Obligatorische Beschreibung**

Die Felder in diesem Block bilden die minimal notwendigen Eigenschaften der Sammlungsbeschreibung und sind Pflichtfelder. Eine Beschreibung der Sammlung muss in mindestens einer Sprache vorliegen.

Beschreibungen	Titel	Acronym	Sprache
	Bayerisches Digitales Repository	BDR	deu

+ Beschreibung hinzufügen

Contains the basic descriptive elements 'title', 'acronym' and 'description' Each description set is associated with a specific language allowing multilingual collection descriptions. Please note, that the first entry is considered as primary language for the collection and serves as displayed default.

Kollektionstyp: Digitale Sammlung

**Akteur gespeichert**

Laczny Joachim

5. Dezember 2016 10:19:46  
hannastolz@dariah.eu

**Versionshistorie**

- 28. Oktober 2016 14:53:16 (Aktuellste)
  - Veröffentlicht
  - tobiasgradl@dariah.eu
- 28. Oktober 2016 14:53:09
  - Entwurf
  - tobiasgradl@dariah.eu

© DARIAH-DE Impressum Kontakt

Source: <https://colreg.de.dariah.eu>

~ DARIAH-DE    DARIAH-EU    DHd-Blog    DHd-Kanal    Twitter    TextGrid    [Sign Out](#)

## DARIAH-DE Publikator


On this page you can edit your chosen collection. Please **add files** from your disc drive and add metadata to your collection and files. As metadata schema Dublin Core will be used, to provide a minimal standard and a certain flexibility for different research data formats. For that reason only few of the metadata fields are mandatory. Please note: the more metadata you provide with your collection and files, the bigger the chance that your data (and of course metadata) is found and re-used by other researchers and can be put into the research data life cycle again.

### Edit Collection

[← back to main view](#)

– Manage Content and Edit Metadata

[add file\(s\)](#)      [show optional metadata](#)

 [view raw collection data](#)

\* **Title** ⓘ ⬆

Please provide mandatory metadata

\* **Creator** ⓘ ⬆

Please provide mandatory metadata

\* **Rights** ⓘ ⬆

Please provide mandatory metadata

Alle Inhalte dieser Seite unterliegen sofern nicht anders gekennzeichnet der CC-BY 4.0 Lizenz      [Impressum](#)    [Kontakt](#)

Source: <https://devportal.de.dariah.eu/web/dariah-imp1/publish>

- Verzeichnis von Forschungsdatensammlungen
  - relevant für **Forschungsfrage / einzelne Benutzer**
  - **strukturierte und semi-strukturierte** Inhalte
  - keine grundsätzlichen **Qualitätsvorgaben**
- Definition und Beschreibung von **Datenmodellen**
- **Assoziation** von Daten zur Vorbereitung
  - **Übergreifender Suchen** in den Inhalten heterogener Sammlungen
  - **Forschungsorientierter Sichten** auf Daten
  - **Vergleiche** von Kollektionen und Inhalten



## DARIAH-DE Repository Publish GUI

<https://de.dariah.eu/repository>

## Collection Registry

<http://demo2.dariah.eu/colreg/>

## Generic Search

<http://dev3.dariah.eu/search/>

## TextGrid Repository

<https://textgridrep.org/en/>

Herzlichen Dank für  
Ihre Aufmerksamkeit

Tobias Gradl

[tobias.gradl@uni-bamberg.de](mailto:tobias.gradl@uni-bamberg.de)

Dr. Stefan Schmunk

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

 @DARIAHde