



# DISCUSS DATA

Discuss Data und Community Spaces. Community-zentriertes  
Forschungsdatenmanagement für die Geistes- und Sozialwissenschaften

Torsten Kahlert (HAB Wolfenbüttel) & Daniel Kurzawe (SUB Göttingen)

Teil I

**Discuss Data: *Überblick***

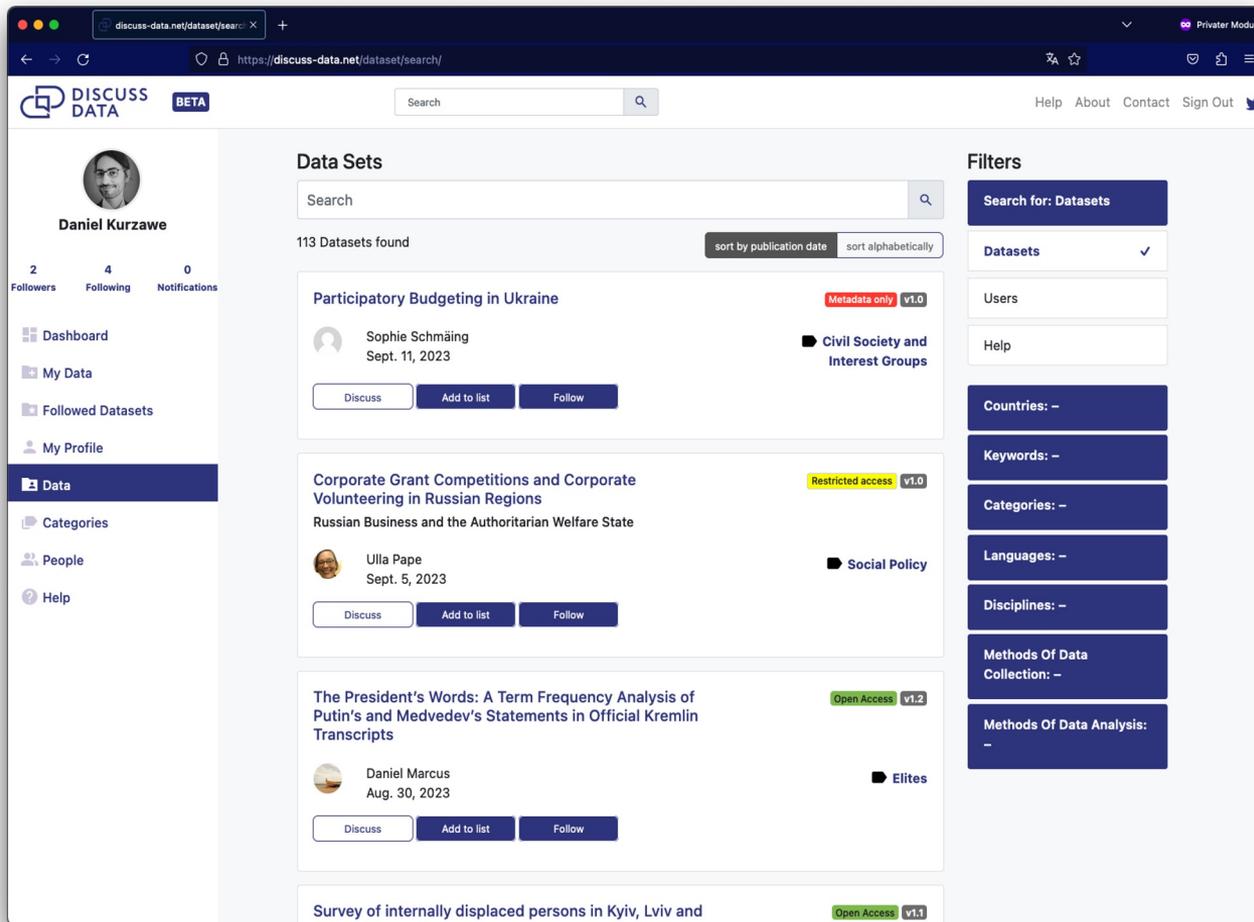
Teil II

**Community Space für die *Digital Humanities***



*Discuss Data I:  
Ein Überblick*

- Discuss Data ist ein **Ort für die Datenpublikation**
- Besonderheit: **Diskussionen zu Daten-Kontexten, -Methoden und -qualität** auf Discuss Data
- **Phase I:** Community zur Regionalforschung zu “Eastern Europe, South Caucasus and Central Asia” – ein sicherer Ort für Daten aus autoritären Kontexten
- **Phase II:** Modularisierung und zusätzliche Spaces für Digital Humanities und Archivdaten
- Discuss Data setzt auf **FAIR Data**, erlaubt auch zusätzlich ein umfangreiches Rechtemanagement und fachspezifische Lizenzen
- Eingebettet in **Infrastrukturen:** DARIAH-DE, NFDI Text+ und 4Memory, Base4NFDI



The screenshot shows a web browser window with the URL `https://discuss-data.net/dataset/search/`. The page features a navigation bar with the site logo, a search bar, and links for Help, About, Contact, and Sign Out. A user profile for Daniel Kurzawe is visible on the left, with 2 followers, 4 following, and 0 notifications. The main content area is titled "Data Sets" and shows 113 datasets found, sorted by publication date. Three datasets are visible:

- Participatory Budgeting in Ukraine** by Sophie Schmäing, Sept. 11, 2023. Metadata only v1.0. Category: Civil Society and Interest Groups.
- Corporate Grant Competitions and Corporate Volunteering in Russian Regions** by Ulla Pape, Sept. 5, 2023. Restricted access v1.0. Category: Social Policy.
- The President's Words: A Term Frequency Analysis of Putin's and Medvedev's Statements in Official Kremlin Transcripts** by Daniel Marcus, Aug. 30, 2023. Open Access v1.2. Category: Elites.

A right-hand sidebar contains a "Filters" section with various dropdown menus: Search for: Datasets, Datasets (checked), Users, Help, Countries: -, Keywords: -, Categories: -, Languages: -, Disciplines: -, Methods Of Data Collection: -, and Methods Of Data Analysis: -.

- Gefördert seit 2016 durch DFG, seit 2020 in offener Beta online
- Open Source Architektur in aktiver Entwicklung (2187 commits, 220 releases)
- DOIs für Referenzierung
- Publikation über DARIAH-DE Repository (weitere Optionen werden im Zuge der Modularisierung implementiert)

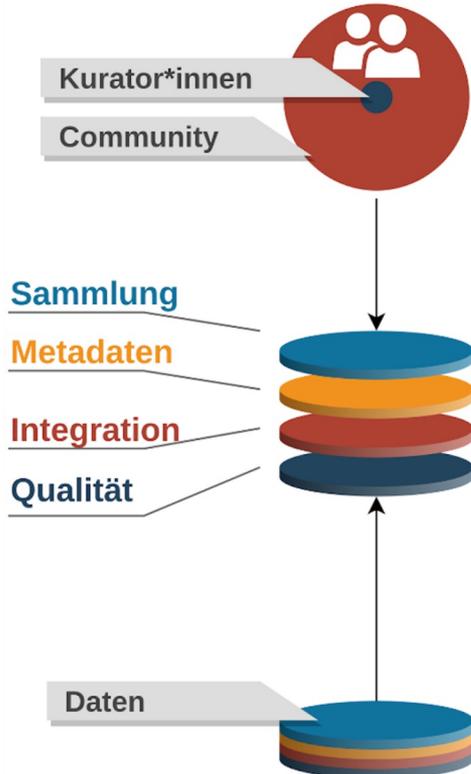
## *Community Space Eastern Europe, South Caucasus and Central Asia*

- 161 Accounts (über DARIAH-DE AAI Infrastruktur, weitere Wege sind geplant)
- 147 Datensätze (Open Access und Restricted Access)
- 16 Kategorien
- 13 Kurator\*innen

Community und Daten



Services



Evaluation und  
Diskussion

Vernetzung

Dokumentation

Datenhaltung



Persistente  
Identifizierung



Authentifizierung  
Autorisierung

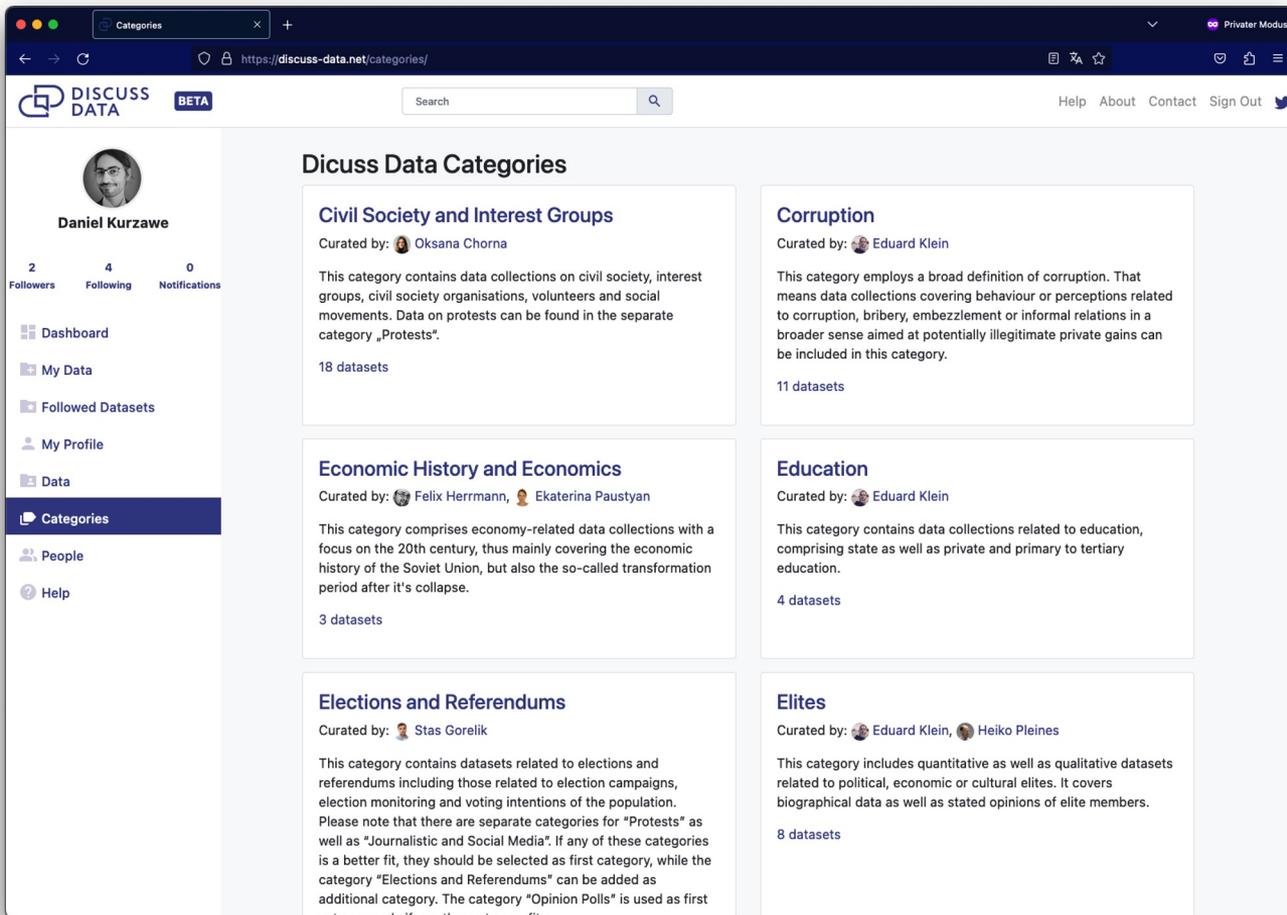


Persistente  
Speicherung



Hosting, Storage  
und Backup

- Wir verstehen unter Forschungscommunities:
  - übergreifende und disziplin unabhängige Gruppen
  - Definieren sich über Schnittmengen in verwendeten Methoden, Fragestellungen, Vokabularien und Daten(typen)
- Ziel des Community-zentrierten Ansatzes:
  - stärkere Einbindung von Forscher\*innen in FDM Infrastrukturen
  - Schaffung eines Kommunikationsraumes für (Sub-)Communities um Daten aus ihrem Bereich teilen, kuratieren und diskutieren zu können
- Im Fall der interdisziplinär verfassten area studies zum post-sowjetischen Raum
  - sind das (Sub-)Communities z.B. zu Protesten, Korruption, Sozialpolitik, Bildung, Wirtschaft
  - aus den Fachbereichen Geschichtswissenschaften, Politikwissenschaften, Soziologie, Kulturwissenschaften, Ethnologie, ...



The screenshot shows a web browser window displaying the 'Dicuss Data Categories' page on the DISCUSS-DATA.NET website. The browser's address bar shows the URL <https://discuss-data.net/categories/>. The website header includes the logo, a search bar, and navigation links for Help, About, Contact, and Sign Out. A user profile for Daniel Kurzawe is visible on the left sidebar, with 2 followers, 4 following, and 0 notifications. The main content area is titled 'Dicuss Data Categories' and lists six categories, each with a curator, a description, and the number of datasets.

## Dicuss Data Categories

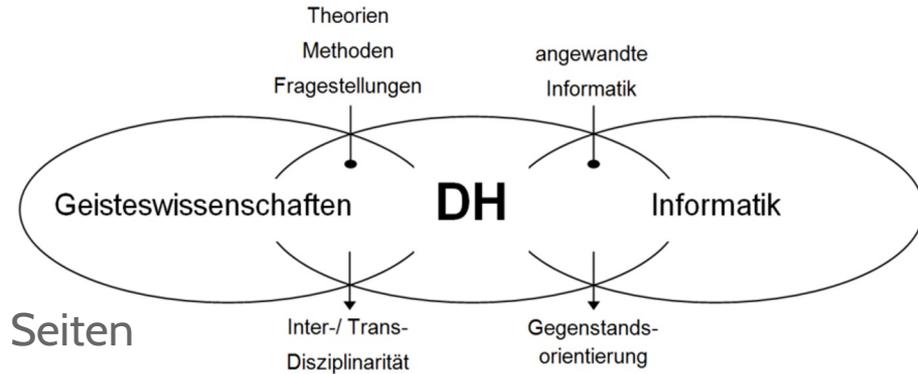
- Civil Society and Interest Groups**  
Curated by:  Oksana Chorna  
This category contains data collections on civil society, interest groups, civil society organisations, volunteers and social movements. Data on protests can be found in the separate category „Protests“.  
18 datasets
- Corruption**  
Curated by:  Eduard Klein  
This category employs a broad definition of corruption. That means data collections covering behaviour or perceptions related to corruption, bribery, embezzlement or informal relations in a broader sense aimed at potentially illegitimate private gains can be included in this category.  
11 datasets
- Economic History and Economics**  
Curated by:  Felix Herrmann,  Ekaterina Paustyan  
This category comprises economy-related data collections with a focus on the 20th century, thus mainly covering the economic history of the Soviet Union, but also the so-called transformation period after it's collapse.  
3 datasets
- Education**  
Curated by:  Eduard Klein  
This category contains data collections related to education, comprising state as well as private and primary to tertiary education.  
4 datasets
- Elections and Referendums**  
Curated by:  Stas Gorelik  
This category contains datasets related to elections and referendums including those related to election campaigns, election monitoring and voting intentions of the population. Please note that there are separate categories for “Protests” as well as “Journalistic and Social Media”. If any of these categories is a better fit, they should be selected as first category, while the category “Elections and Referendums” can be added as additional category. The category “Opinion Polls” is used as first category only if no other category fits.
- Elites**  
Curated by:  Eduard Klein,  Heiko Pleines  
This category includes quantitative as well as qualitative datasets related to political, economic or cultural elites. It covers biographical data as well as stated opinions of elite members.  
8 datasets

- Aktuelle Community sieht Bedarfe und nutzt Publikationen
- Großes Interesse aus vielen unterschiedlichen Forschungsbereichen
- Erfahrungen mit dem Kernkonzept: zur Verbindung von Diskurs und Daten
  - Diskussion eher spärlich und bedarfen Motivation → Hemmschwelle?
  - Aber: Funktion wird immer wieder als positiv bewertet → Andere Repräsentation notwendig?
    - Design: Diskussionsfunktion zu versteckt
    - Kultur: Diskutiert auf Konferenzen und mündlich oder auf sozialen Medien schriftlich, jedoch nicht zu Daten in Repositorien
    - Aufwand im zu erwartenden Ertrag noch zu gering?
- neue Lösungsansätze:
  - Zusammenarbeit mit einschlägigen Fachzeitschriften für gemeinsame Erarbeitung und Publikation von Reviews zu Forschungsdaten oder auch Forschungsdatenreviews (die Forschungsdaten mehrerer Repositorien betrachten)



# Discuss Data II: Community Space für die *Digital Humanities*

- **Digital Humanities** an der Schnittstelle von Informatik & Geisteswissenschaften
- Methoden und Arbeitsweisen von beiden Seiten
- DH driving force für Open Science und FAIR data
- als **data driven research** spielen **Forschungsdaten** eine prominente Rolle
- es gibt Bedarf an Möglichkeiten für kuratierte Diskussion über FD
- DH Community mit **digitalen öffentlichen Räumen** vertrauter als traditionsreichere Fachcommunities

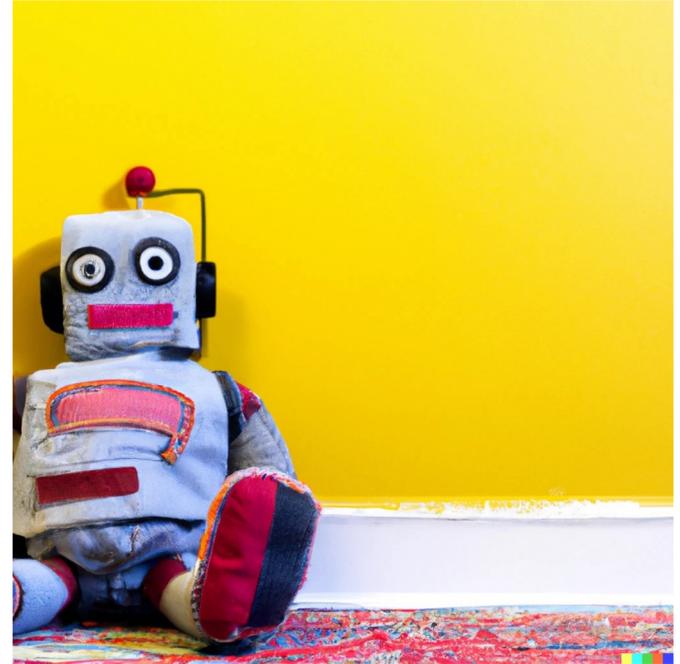


Quelle: <https://www.geschichte.uni-wuppertal.de/de/lehrgebiete/digital-humanities/>

- Aufbau einer Redaktion von **Datenkurator\*innen**
- Förderung von **Data Literacy**
- Weiterentwicklung von **Datenstandards** als Empfehlungen für die Community
- Aufbau von neuen Formaten wie **Data Reviews**
- Neugestaltung der **Diskussionsfunktion**
- Einbindung in Infrastrukturen von **PIDs** (DOIs, ORCID, ROR etc.)
- Anbindung an weitere Repositorien (z.B. Zenodo, github)
- **Kooperationen** mit DH-Zentren und Forschungsinstituten

- geisteswissenschaftliche Datenproduktion **verstehen**
- “sharing is caring” (digitale) Geisteswissenschaftler\*innen davon zu **überzeugen**, ihre Daten zu teilen
- Beitrag leisten zur **Verbesserung der Reputation** von Datenpublikationen in der Scholarly Community
- **Diskussion** auf Discuss Data stärker in Gang bringen, über neue Formate wie Diskussionsbeiträge als **Mikropublikationen und bessere Sichtbarkeit**

- Umgang mit lizenz- oder urheberrechtlich sensiblen **Materialien**
- **Modularisierung** der Plattform
- Nachhaltigkeit: **langfristige Perspektive** für das Community-Modell und die Selbstverwaltung  
Discuss Data



Quelle: dall.e



*Vielen Dank!*