



# formemory

Task Area 2: Data Connectivity (4Memory)  
“Register für historische und  
objektbezogene Vokabulare und Normdaten”







LimeSurvey-Testversion

<https://umfrage.uni-halle.de/458285?lang=de>

LimeSurvey-Registererfassung

<https://umfrage.uni-halle.de/194457?lang=de>

FactGrid

<https://database.factgrid.de/wiki/Item:Q518220>



TA2

Interviews  
Use Cases

Register  
kontrollierte  
Vokabulare

Empfehlungen  
Guidelines

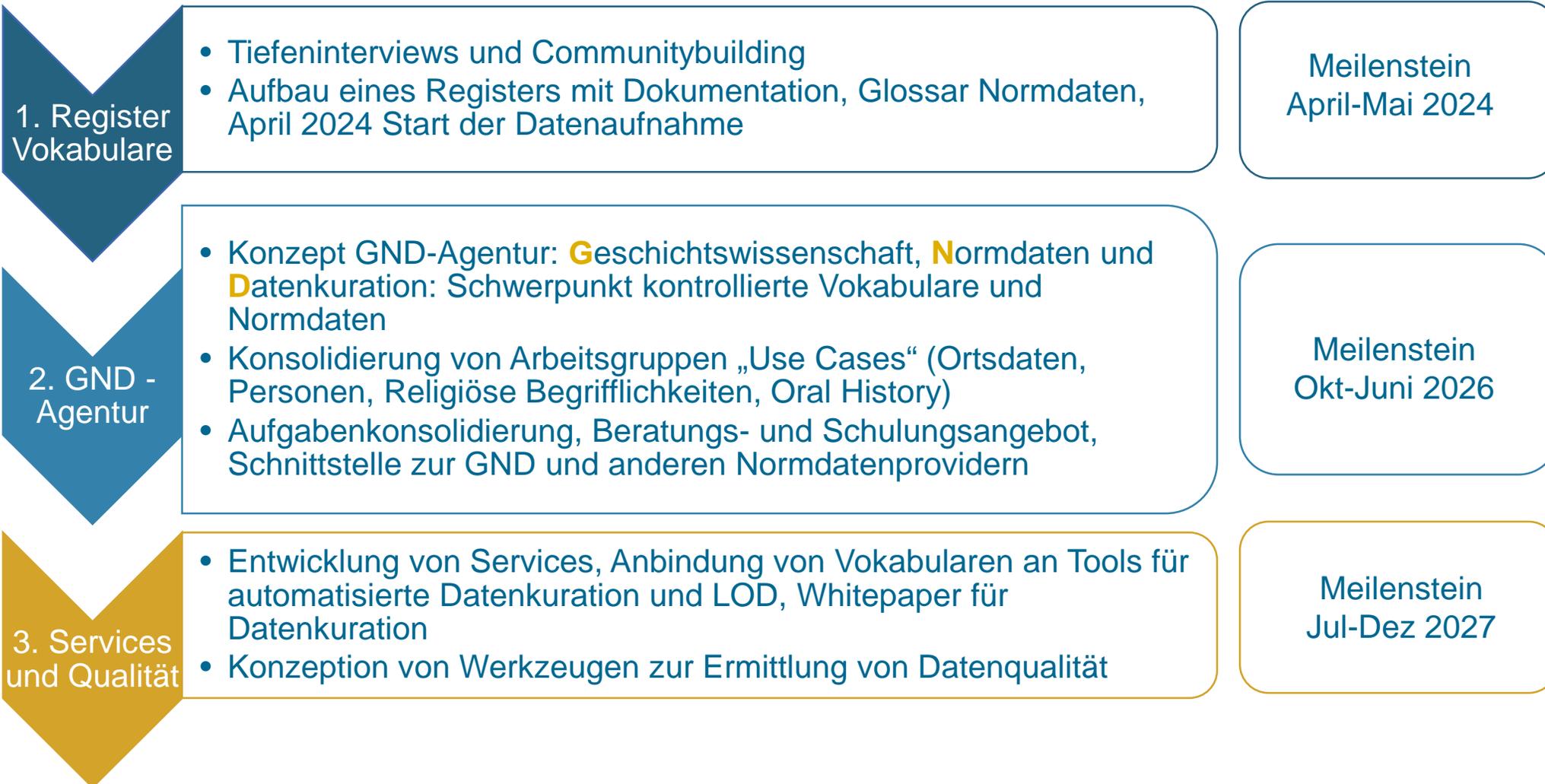
Beratung  
und Services  
Daten-  
kuration

Tools und  
Werkzeuge

GND  
Agentur

[ta2-nfdi4memory@geschichte.uni-halle.de](mailto:ta2-nfdi4memory@geschichte.uni-halle.de)

Bild: Katja Dohnke



- NFDI-Konsortium: NFDI4Memory
- Task Area 2 „Data Connectivity“
- E-Mail [[ta2-nfdi4memory@geschichte.uni-halle.de](mailto:ta2-nfdi4memory@geschichte.uni-halle.de)]
- Mailingliste der TA2: Data Connectivity [[nfdi4memory-ta2@lists.nfdi.de](mailto:nfdi4memory-ta2@lists.nfdi.de)]
- Blog der Task Area Data Connectivity [<https://blogs.urz.uni-halle.de/nfdi4memory/>]
- Newsletter von NFDI4Memory [<https://4memory.de/>]
- NFDI-Konsortium: NFDI4Objects
- Task Area 6 “Qualification, Integration and Harmonisation”
- E-Mail des Helpdesks [n4o-helpdesk@dainst.de](mailto:n4o-helpdesk@dainst.de)
- <https://www.nfdi4objects.net/>
- Newsletter <https://www.listserv.dfn.de/sympa/info/nfdi4objects>.



- Register für historische und objektbasierte Vokabulare und Normdaten
- Verzeichnis über relevante Datenquellen und Möglichkeiten und Bedingungen ihrer Nutzung
- Instrument und Grundlage zur Datenvernetzung und zum Matching
- Bessere Nutzungsmöglichkeiten und Zugänglichkeit zu Ressourcen und Quellen



- Übersicht zu vorhandenen Angeboten und Vokabularen
  - BARTOC
  - Museumsverband, AG mit Mitwirkungsmöglichkeiten
  - DigiCult
  - Einzelprojekte
  - kostenpflichtige Projekte
  - Internationale Standards, Initiativen und Angebote
- Verbesserung der Informationstiefe und Findbarkeit
- Zugänglichkeit und Anbindbarkeit
- Einfachere, schnellere Nutzbarkeit und Mitwirkungsmöglichkeiten

- FAIR-Prinzipien und Möglichkeiten der Recherche von Daten bis hin zur Verbesserung von Volltexterkennung
- Data Linkage: Datenmatching, Dubletten- bzw. Identitätserkennung, webbasiertes Linkage (LOD)
- Datenkuration und Qualitätskontrolle: Datenbereinigung, Plausibilitätstests und Qualitätsprüfung
- Anreicherung und Analyse: Entitätserkennung mittels KI-Technologien, Hierarchische Vokabularien zur Kategorisierung und Analyse von Daten



Bild: Dr. Marcus Gossler, Schlagwortkatalog 1926–1993 der Universitätsbibliothek Graz, CC BY-SA 3.0

## Entität

- Klasse von Begriffen oder Kategorien wie Orte, Berufe, Personen, die aufgrund bestimmter Definitionen gemeinsame Eigenschaften besitzen

## Wortlisten und Kategoriensysteme

- Zugeordnete Definitionen, Konzepte oder Beschreibungen
- Dokumentationen
- Überindividuelle Nachvollziehbarkeit von Analyseprozessen

## Thesaurus / Thesauri

- Regelbasierte Normierung natürlicher Sprache und semantischer Formen/Inhalte
- Relation und Verknüpfung von Wörtern ähnlicher oder gegensätzlicher Bedeutung
- Weniger Klassifikation, sondern assoziative, nichthierarchische Relationen

## Klassifikationen / Taxonomien

- Messen aufgrund einer Methode bestimmte Eigenschaften einer Entität
- Organisieren darauf aufbauend eine hierarchische Struktur der Informationen
- Regelbasiertes System

## Normdaten

- Öffentlich verfügbare, standardisierte Thesauri oder Klassifikationen die durch eine größere Community gemeinsam genutzt und von einem bestimmten Personenkreis gepflegt werden sowie über Identifikatoren und eine dauerhafte Organisation persistent bleiben

## Ontologie / Schemata

- Informatisches Konzept, um Vokabulare in technische Schemata, formale Modelle und webbasierte Protokolle zu überführen. Sie beschreiben die Beziehung zwischen unterschiedlichen Begriffen, Konzepten und Ordnungsprinzipien

- Regelbasiertes Management der Arbeit mit Vokabularen und Normdaten entwickeln (abgestuftes Konzept) – Instanz für Linked Open Data zwischen Einzelprojekt und Gemeinsamer Normdatei (GND) = FactGRID, DigiCult, Dante
- Möglichkeit zur Nutzung der Wikibase-Instanz FactGRID, Beratungs- und Schulungsmöglichkeiten

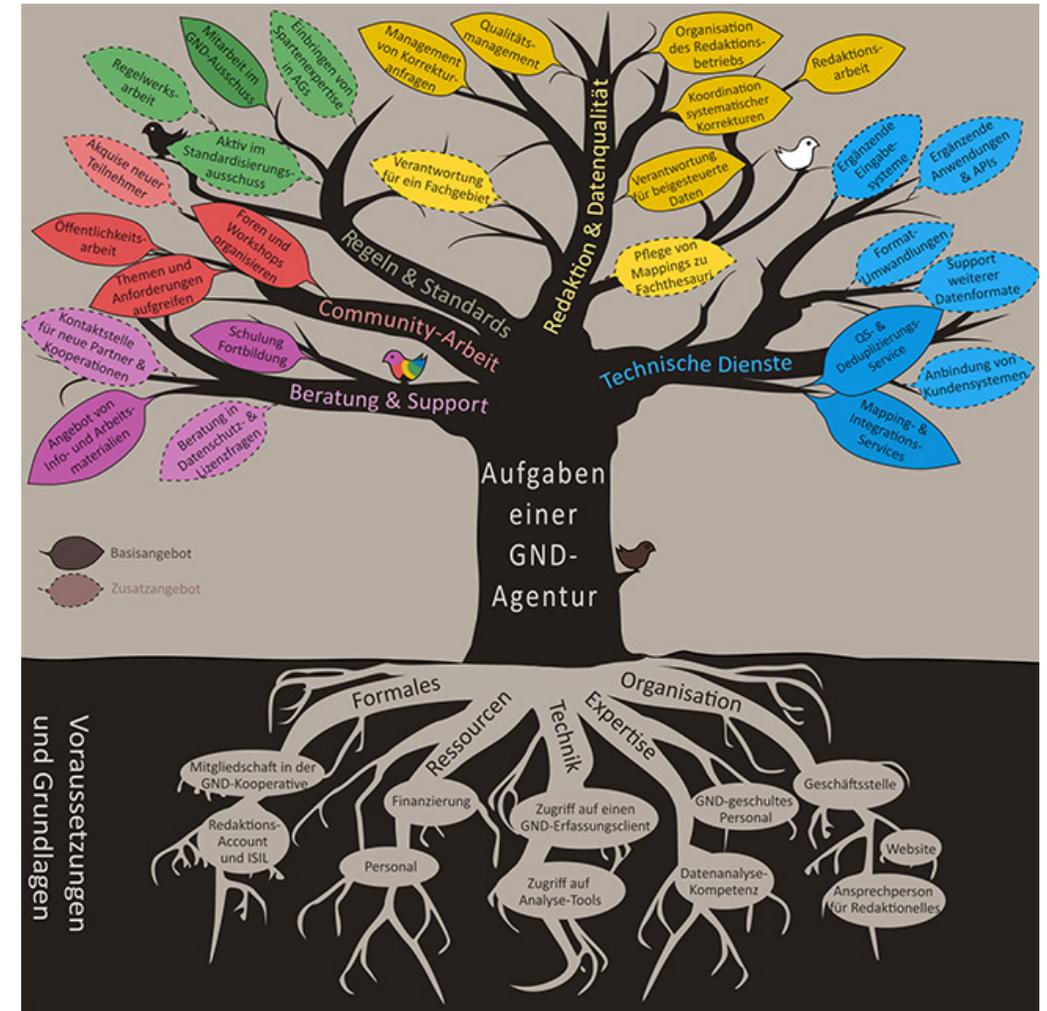
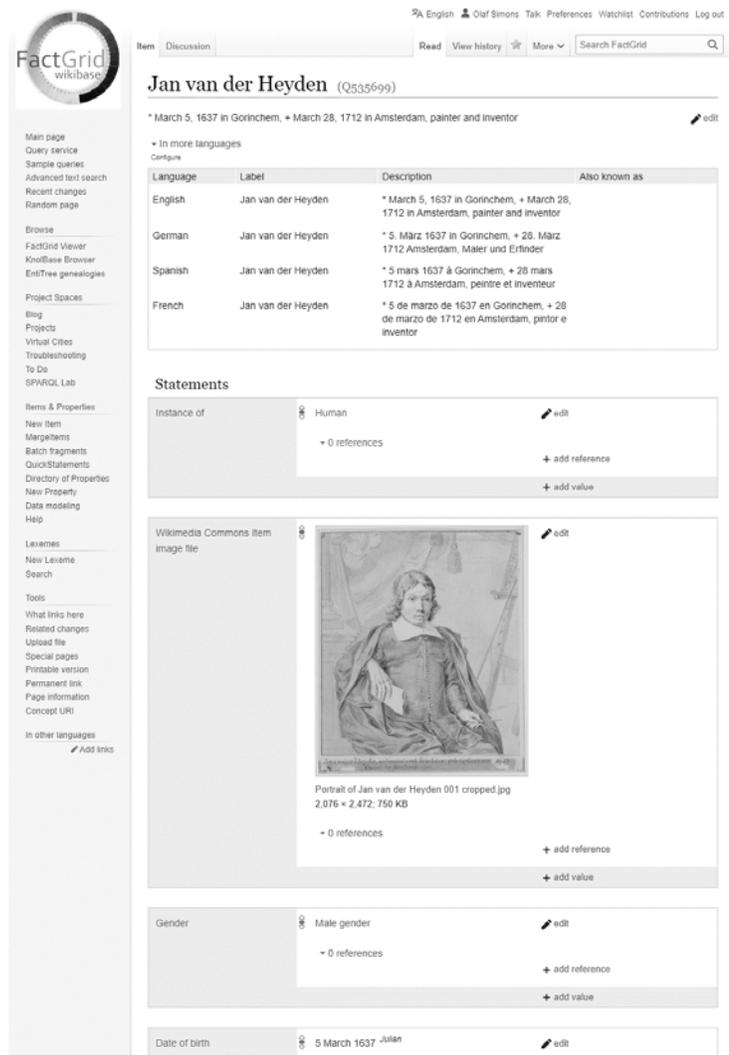


Bild: J. Kett (DNB) CC BY SA



<https://database.factgrid.de>



The screenshot shows the Wikibase interface for the entry 'Jan van der Heyden' (Q535699). The main content area displays a table with columns for Language, Label, Description, and Also known as. The table contains entries for English, German, Spanish, and French. Below the table, there are sections for Statements, Wikidata Commons item image file, Gender, and Date of birth. The interface includes a sidebar with navigation options and a top navigation bar with user information and search options.

Language	Label	Description	Also known as
English	Jan van der Heyden	* March 5, 1637 in Gorinchem, + March 28, 1712 in Amsterdam, painter and inventor	
German	Jan van der Heyden	* 5. März 1637 in Gorinchem, + 28. März 1712 Amsterdam, Maler und Erfinder	
Spanish	Jan van der Heyden	* 5 mars 1637 à Gorinchem, + 28 mars 1712 à Amsterdam, peintre et inventeur	
French	Jan van der Heyden	* 5 de marzo de 1637 en Gorinchem, + 28 de marzo de 1712 en Amsterdam, pintor e inventor	

## FactGrid

- Wikibase Graph-Datenbank (RDF)
- 400 registrierte Nutzer
- ca. 50 aktive Projekte von Alt-Orientalistik bis Gegenwart
- 900.000 Datenbankobjekte
- derzeit mit Schwerpunkten Deutschland, Frankreich, Spanien, Ungarn, USA – technisch jedoch beliebig mehrsprachig (durchgehend Deutsch, Englisch, Französisch, Spanisch)
- Daten CC0-lizenziert über SPARQL-Endpoint herunterladbar

## Zielsetzung:

- nachhaltiges Datenmanagement: Daten laufen unabhängig von Projektlaufzeiten und Finanzierungen auf einer kollaborativ gepflegten Plattform
- globale Vernetzung von Projekten
- ein Datenrepositorium – beliebig viele lokale Projektfrontends



## Dokumentation zum Register für historische und objektorientierte Vokabulare und Normdaten (R:hovono)

Julian Freytag, <https://orcid.org/0009-0002-0622-2184>

Katja Liebing, <https://orcid.org/0009-0001-1624-6465>

Katrin Moeller, <https://orcid.org/0000-0003-4090-5667>

Anne Purschwitz, <https://orcid.org/0000-0002-2754-8792>

Olaf Simons, <https://orcid.org/0000-0001-9230-4666>

Marius Wegener, <https://orcid.org/0009-0007-6782-1865>

Kontakt: TA2 NFDI4Memory <ta2-nfdi4memory@geschichte.uni-halle.de>  
Eine Arbeit im Rahmen des 4Memory-Konsortiums.<sup>7</sup>

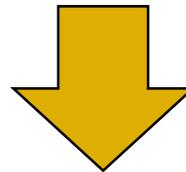
- Dokumentation/Codebuch/Glossar noch in Bearbeitung, werden in Kürze bei Zenodo veröffentlicht werden
- Zugriff auf Datenpaper:
  - Julian Freytag, Katja Liebing, Katrin Moeller, Anne Purschwitz, Olaf Simon und Marius Wegener: Codebuch zur LimeSurvey-Umfrage und FactGrid-Eintragung für das Register historischer und objektbezogener Vokabulare und Normdaten (R:hovono) / Codebook for the LimeSurvey survey and FactGrid entry for the register of historical and object-related vocabularies and authority data (R:hovono), Halle 2024, DOI: 10.5281/zenodo.11031743. <https://zenodo.org/records/11031743>
  - Julian Freytag, Katja Liebing, Katrin Moeller, Anne Purschwitz, Olaf Simon und Marius Wegener: Dokumentation zum Register historischer und objektorientierter Vokabulare und Normdaten und Vokabulare (R:hovono) / Documentation on the register of historical and object-orientated vocabularies and authority data and vocabularies (R:hovono), Halle 2024, DOI: 10.5281/zenodo.11033367. <https://zenodo.org/records/11033367>

1. Einleitung: Ausgangssituation und Zielsetzung
2. Beteiligte
3. Zielgruppe
4. Bedarfsermittlung: Welches Register braucht die historische Community?
5. Der Erfassungsbogen – Wie werden die Vokabulare für das Register erhoben?
  - 5.1 Erfassungsmethoden
  - 5.2 Erfassung über LimeSurvey - Aufbau und Gliederung des Erfassungsbogens
    - 5.2.1 Kopfdaten/Metadaten
    - 5.2.2 Zugänglichkeit/Lizenzen
    - 5.2.3 Crosskonkordanz
    - 5.2.4 Bereitstellung der Daten
    - 5.2.5 Ausblick und optionale Angaben
  - 5.3. Einpflegen in FactGrid
- 6 Ausblick / Zeitplan

- Register wird von NFDI4Memory Arbeitsgruppe Data Connectivity am Historischen Datenzentrum Sachsen-Anhalt angeboten und gepflegt
- Auch NFDI4Objects wird R:hovono zur Erfassung und Bereitstellung von Vokabularen verwenden

- Participants
- Alle, die über einschlägige Vokabulare verfügen und diese der Forschung zur Verfügung stellen möchten

- in Antragsphase: 95 Problem Stories
- Umfrage Arbeitsweisen, Erschließungsprinzipien, Nutzungsmöglichkeiten von Vokabulären
- Tiefeninterviews
- NFDI Community Forum



Zielvorstellungen für „Register historischer und objektbezogener Vokabulare und Normdaten“

- LimeSurvey: Programm zu Erstellung und Veröffentlichung von Online-Umfragen
  - „Klassische“ Umfrage → einfache Zugänglichkeit; stellt sicher, dass erforderliche Angaben ausgefüllt
  - Ermöglicht selektive Abfragen und strukturierte Ausgabe
  - (semi)automatische Übertragung in FactGrid möglich
- FactGrid
  - Basiert auf Wikibase, spricht speziell wissenschaftliche Community historisch arbeitender Disziplinen an
  - Erfassungsmethode für erfahrene Nutzer\*innen
  - Einrichtung Nutzerkonto per Mail an: [ta2-nfdi4memory@geschichte.uni-halle.de](mailto:ta2-nfdi4memory@geschichte.uni-halle.de)
- im Entstehen begriffene Website
  - soll strukturiertes Ausgabeformat ermöglichen, um auch niedrighschwelligen Zugang zu Arbeitsergebnissen zu ermöglichen

- 48 Fragen in Bereichen Metadaten, Zugänglichkeit, Crosskondordanzen, Bereitstellung der Daten (Lizenzen) und Ausblick
- Kerndatenset (Titel, Beschreibung, Schlagwörter...) verpflichtend wegen Crosskondanz zu BARTOC (Erkennungszeichen !)
- Je nach Zweck offene Fragen mit Freifeldern, Multiple Choice oder Kombination offener und geschlossener Frage; Filterfragen
- Ermöglicht Übernahme von Nachweis des Vokabulars in Register

- Allgemeine formale und inhaltliche Erschließung, beispielsweise:
  - Kontaktinformationen
  - Titel
  - zeitliche Einordnung
  - Umfang
  - Sprachen
  - inhaltliche Beschreibung
  - Schlagwörter
  - Quellengrundlage

- Erhebung Arten/Bedingungen des Zugangs, relevant für Vernetzung und Übernahme durch Forschende, beispielsweise:
  - Art des Zugangs
  - Lizenzierung
  - Versionierung
  - Aktualisierung und Dokumentation
  - Vollständigkeit
  - Bearbeitungs- und Ergänzungsoptionen
  - Erschließungsgrundsätze

- Verbindungen eines Vokabulars zu anderen Vokabularen auf Ebene von Begriffen oder übergeordneten Kategorien, beispielsweise:
  - Matchings (Erleichterung Datenkuration/Datenanreicherung)
  - Geben Eindruck, mit welchen anderen Vokabularen abgeglichen
  - Erhebung von Nutzergruppen (für Vorschlagsoption Präsentation Webseite)

- Technische Standards, beispielsweise:
  - Ein- und Ausgabeformate
  - API-Schnittstelle
  - SPARQL Endpunkte
  - bestehende Services zum Vokabular

- Bearbeitungsstand des Vokabulars
- Fehlende Erfassung

- Heute: Freischaltung Umfrage
- Erfassung in gesamtem Förderzeitraum möglich, fortwährendes Einpflegen
- Aus Erfahrungen von Nutzenden und Projektverantwortlicher bei Erfassung → Evaluation von Besonderheiten, Rückmeldungen, fortwährende Optimierung des Erfassungsbogens
- ½-jährliche Erinnerung an Einpflegen von Veränderungen
- Quartalsweise Übernahme Daten in FactGrid, Meldung bei BARTOC

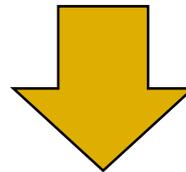
- Veröffentlichung Dokumentation, Codebuch, Glossar
- FactGrid: Individuelle Schulungen

1. Einleitung: Ausgangssituation und Zielsetzung
2. Beteiligte
3. Zielgruppe
4. Bedarfsermittlung: Welches Register braucht die historische Community?
5. Der Erfassungsbogen – Wie werden die Vokabulare für das Register erhoben?
  - 5.1 Erfassungsmethoden
  - 5.2 Erfassung über LimeSurvey - Aufbau und Gliederung des Erfassungsbogens
    - 5.2.1 Kopfdaten/Metadaten
    - 5.2.2 Zugänglichkeit/Lizenzen
    - 5.2.3 Crosskonkordanz
    - 5.2.4 Bereitstellung der Daten
    - 5.2.5 Ausblick und optionale Angaben
  - 5.3. Einpflegen in FactGrid
- 6 Ausblick / Zeitplan

- Register wird von NFDI4Memory Arbeitsgruppe Data Connectivity am Historischen Datenzentrum Sachsen-Anhalt angeboten und gepflegt
- Auch NFDI4Objects wird R:hovono zur Erfassung und Bereitstellung von Vokabularen verwenden

- Participants
- Alle, die über einschlägige Vokabulare verfügen und diese der Forschung zur Verfügung stellen möchten

- in Antragsphase: 95 Problem Stories
- Umfrage Arbeitsweisen, Erschließungsprinzipien, Nutzungsmöglichkeiten von Vokabulären
- Tiefeninterviews
- NFDI Community Forum



Zielvorstellungen für „Register historischer und objektbezogener Vokabulare und Normdaten“

- Lime Survey: Programm zu Erstellung und Veröffentlichung von Online-Umfragen
  - „Klassische“ Umfrage → einfache Zugänglichkeit; stellt sicher, dass erforderliche Angaben ausgefüllt
  - Ermöglicht selektive Abfragen und strukturierte Ausgabe
  - (semi)automatische Übertragung in FactGrid möglich
- FactGrid
  - Basiert auf Wikibase, spricht speziell wissenschaftliche Community historisch arbeitender Disziplinen an
  - Erfassungsmethode für erfahrene Nutzer\*innen
  - Einrichtung Nutzerkonto per Mail an: [ta2-nfdi4memory@geschichte.uni-halle.de](mailto:ta2-nfdi4memory@geschichte.uni-halle.de)
- im Entstehen begriffene Website
  - soll strukturiertes Ausgabeformat ermöglichen, um auch niedrighschwelligen Zugang zu Arbeitsergebnissen zu ermöglichen

- 48 Fragen in Bereichen Metadaten, Zugänglichkeit, Crosskondordanzen, Bereitstellung der Daten (Lizenzen) und Ausblick
- Kerndatenset (Titel, Beschreibung, Schlagwörter...) verpflichtend wegen Crosskondanz zu BARTOC (Erkennungszeichen !)
- Je nach Zweck offene Fragen mit Freifeldern, Multiple Choice oder Kombination offener und geschlossener Frage; Filterfragen
- Ermöglicht Übernahme von Nachweis des Vokabulars in Register

- Allgemeine formale und inhaltliche Erschließung, beispielsweise:
  - Kontaktinformationen
  - Titel
  - zeitliche Einordnung
  - Umfang
  - Sprachen
  - inhaltliche Beschreibung
  - Schlagwörter
  - Quellengrundlage

- Erhebung Arten/Bedingungen des Zugangs, relevant für Vernetzung und Übernahme durch Forschende, beispielsweise:
  - Art des Zugangs
  - Lizenzierung
  - Versionierung
  - Aktualisierung und Dokumentation
  - Vollständigkeit
  - Bearbeitungs- und Ergänzungsoptionen
  - Erschließungsgrundsätze

- Verbindungen eines Vokabulars zu anderen Vokabularen auf Ebene von Begriffen oder übergeordneten Kategorien, beispielsweise:
  - Matchings (Erleichterung Datenkuration/Datenanreicherung)
  - Geben Eindruck, mit welchen anderen Vokabularen abgeglichen
  - Erhebung von Nutzergruppen (für Vorschlagsoption Präsentation Webseite)

- Technische Standards, beispielsweise:
  - Ein- und Ausgabeformate
  - API-Schnittstelle
  - SPARQL Endpunkte
  - bestehende Services zum Vokabular

- Bearbeitungsstand des Vokabulars
- Fehlende Erfassung

- Heute: Freischaltung Umfrage
- Erfassung in gesamtem Förderzeitraum möglich, fortwährendes Einpflegen
- Aus Erfahrungen von Nutzenden und Projektverantwortlicher bei Erfassung → Evaluation von Besonderheiten, Rückmeldungen, fortwährende Optimierung des Erfassungsbogens
- ½-jährliche Erinnerung an Einpflegen von Veränderungen
- Quartalsweise Übernahme Daten in FactGrid, Meldung bei BARTOC

- Veröffentlichung Dokumentation, Codebuch, Glossar
- FactGrid: Individuelle Schulungen



LimeSurvey

<https://umfrage.uni-halle.de/194457?lang=de>