

Codebuch zur LimeSurvey-Umfrage und FactGrid-Eintragung für das Register historischer und objektbezogener Vokabulare und Normdaten (R:hovono)



Julian Freytag, <https://orcid.org/0009-0002-0622-2184>
Katja Liebing, <https://orcid.org/0009-0001-1624-6465>
Katrín 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.¹

Einleitung



Bei einer Vielzahl von Forschungsprojekten der Geschichtswissenschaften werden inzwischen direkt oder indirekt kontrollierte Vokabulare erstellt. Viele davon sind anderen Forschenden, für die sie nützlich sein könnten, aber nicht bekannt oder auffindbar. Ziel unseres Registers ist es, einen Überblick über die in der historischen Community erstellten Vokabulare und Normdaten zu geben und diese übersichtlich und umfassend darzustellen, um Forschenden Recherche und Nutzung zu erleichtern. Wir beabsichtigen damit einerseits die Sichtbarkeit bereits nachgewiesener oder zugänglicher Vokabulare für die Fachcommunity zu erhöhen und andererseits Vokabulare durch den Austausch mit der Community zu sichten, um einen Nachweis und/oder die Zugänglichkeit sowie auch eine Vernetzung von Vokabularen zu ermöglichen.

In Vorbereitung dessen haben wir uns in Tiefeninterviews mit der Community über Anforderungen an ein solches Register ausgetauscht und unseren Erfassungsbogen nach diesen Kriterien gestaltet. Der daraus hervorgegangene LimeSurvey umfasst die Erhebung von Informationen bezüglich Metadaten, Crosskonkordanzen, der Zugänglichkeit und der Bereitstellung von Daten und dient der Erfassung und Präsentation Ihrer Vokabulare. Sie können mehrere Vokabulare eintragen, nutzen Sie dafür aber bitte für jedes

¹ Diese Arbeit ist im Rahmen des NFDI-Konsortiums 4Memory entstanden (www.4memory.de). Wir danken der Deutschen Forschungsgemeinschaft (DFG) für die finanzielle Unterstützung – Projektnummer 501609550.

ein eigenes Formular. Die im vorliegenden Codebuch enthaltenen Anmerkungen zu den Fragen sollen Ihnen das Ausfüllen des Fragebogens erleichtern.

Sollte es weitere Verständnisfragen oder Verbesserungsvorschläge von Ihrer Seite geben, teilen Sie uns das gerne über die Umfrage, im direkten Gespräch oder per E-Mail unter ta2-nfdi4memory@geschichte.uni-halle.de mit.

Die folgende Darstellung orientiert sich an der Struktur des Erfassungsbogens in LimeSurvey und der dort ab sofort zugänglichen Umfrage:

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

Die im Codebuch abgebildete erste Nummer gibt, die Reihenfolge im Erfassungsbogen wieder, die die Nutzer*innen angezeigt bekommen, dann erfolgt eine kapitelorientierte Nummerierung (in runden Klammern). Der darauf folgende alphanumerische Code, beginnend mit "P", verweist auf das entsprechende Erfassungsfeld in FactGrid.² Er ist lediglich für diejenigen relevant, welche die direkte Eingabe über FactGrid nutzen möchten. Die Struktur des Abfrageschemas in [FactGrid](https://database.factgrid.de/wiki/Item:Q518220) [<https://database.factgrid.de/wiki/Item:Q518220>³] wird in einem separaten Modell ersichtlich. Zudem existieren Filterfragen, deren Antworten Bedingung für Folgefragen sind. Diese besitzen keine P-Nummer, sondern sind mit einem Fragencode nach dem Schema Bed + laufende Nummer (Bed001 usw.) versehen. Danach wurde in eckigen Klammern ein kurzer Verwendungszweck sowie die Frage in LimeSurvey notiert, gefolgt von einer kurzen Erläuterung. Fragen, die hier mit einem Ausrufezeichen gekennzeichnet sind, gehören zum Kernsatz und sind verpflichtend zu beantworten.

Mit dem Ausfüllen der Eingabemaske in LimeSurvey bzw. FactGrid erklären Sie sich damit einverstanden, dass Ihre Angaben im „Register historischer und objektbezogener Vokabulare und Normdaten“ (R:hovono), auf BARTOC und FactGrid sowie in weiteren webbasierten oder analogen Medien veröffentlicht und verbreitet werden. Das Angebot steht insgesamt unter der Lizenz: CC BY 4.0 DEED, Namensnennung 4.0 International, womit Sie sich zu Beginn der Umfrage einverstanden erklären müssen.

Zum Codebuch gehört ebenfalls eine ausführliche Dokumentation der Erhebung des Registers historischer und objektbezogener Vokabulare und Normdaten (R:hovono).⁴

² Die 2018 am Forschungszentrum Gotha in einer Kooperation mit Wikimedia Deutschland gegründete Wikibase-Instanz FactGrid läuft unter der URL <https://database.factgrid.de/>. Mehr Informationen zu ihr bietet der Projektblog <https://blog.factgrid.de/>.

³ Die Informationen des vorliegenden Codebuches für FactGrid lassen sich mit der folgenden Suche tabellarisch abfragen und herunterladen: <https://tinyurl.com/24fcfwce>.

⁴ Julian Freytag, Katja Liebing, Katrin Moeller, Anne Purschwitz, Olaf Simon und Marius Wegener: Dokumentation zum Register historischer und objektbezogener Vokabulare und Normdaten (R:hovono), Halle 2024, DOI: 10.5281/zenodo.11033367.

Teil A: Metadaten

A1⁵ (1/1) P267: [Organisatorischer Kontext] Bei welchem NFDI-Konsortium sind Sie Participant bzw. zu welchem gehört Ihr Vokabular?

Tragen Sie bitte ein, über welches Konsortium Ihr Vokabular aufgenommen wurde.

!A2 (1/2) P34: [Benennung] Bitte geben Sie den Namen Ihres Vokabulars an. Verwenden Sie für jede Sprache ein Textfeld.

Die Auswahlliste der Sprachen führt die am häufigsten gesprochenen Sprachen auf. Die Eingabe weiterer Sprachen kann über das freie Textfeld erfolgen. Orientieren Sie sich hierbei an den Benennungen von ISO 639-2/B

[<https://wiki.dnb.de/download/attachments/90411323/sprachenCodesIso6392.pdf>].

!A3 (1/3) P294: [Akronym] Wie lautet das Akronym Ihres Vokabulars?

Bsp.: Ontologie der historischen, deutschsprachigen Amts- und Berufsbezeichnungen:
Akronym = OhdAB

!A4 (1/4) P121: [Werktyp]: Bitte wählen Sie aus zu welcher Art von Vokabular Ihres gehört.

Wortlisten und Kategoriensysteme: In vielen Projekten entstehen Listen von Bezeichnungen, die bestimmten Definitionen, Konzepten oder Beschreibungen zugeordnet werden, um sprachliche Ausdrücke zu systematisieren. In der Regel werden solche Definitionen und ihre Zuordnungen in Dokumentationen festgehalten. Häufig fungiert die Kategorie gleichzeitig als Identifikator. Diese Kategoriensysteme gelten als kontrollierte Vokabulare, weil sie eine überindividuelle Zuordnung und Nachvollziehbarkeit von Analyseprozessen ermöglichen.

Thesaurus/Thesauri: dienen der regelbasierten Normierung natürlicher Sprache und semantischer Formen/Inhalte, indem sie die Relationen/Verknüpfungen von Wörtern mit ähnlichen oder gegensätzlichen Bedeutungen klären und zuordnen. Thesauri müssen in der Regel nicht im Sinne einer Klassifikation strukturiert sein, sondern besitzen weitgehend assoziative, nicht-hierarchische Relationen. Sie dienen dazu:

- Bezeichnungen (Numerus, Genus, Wortreihenfolge) nach vorgegebenen Regeln zu vereinheitlichen.
- Mehrdeutigkeiten von Bezeichnungen aufzulösen (Disambiguierung) und über feste Mechanismen der terminologischen Kontrolle bestehende Homonyme zuzuordnen.

⁵ Die hier angegebenen Buchstaben vor der Nummerierung erscheinen lediglich in der gedruckten Ausgabe von LimeSurvey. Sie werden hier lediglich zur besseren Orientierung wiedergegeben. In der Onlinefassung der Umfrage erscheinen sie nicht.

- Synonyme und bedeutungsähnliche Bezeichnungen zu identifizieren, Begriffen zuzuordnen und über Vorzugsbezeichnungen (Deskriptoren/Schlagworte) auffindbar zu machen.

Klassifikationen/Taxonomien: werden mithilfe einer Methode zur Organisation von Informationen in hierarchische oder taxonomische Strukturen gebracht, wobei Elemente in Gruppen oder Klassen eingeteilt werden, die auf gemeinsamen Merkmalen und Eigenschaften basieren. Sie sind daher auf eine Entität bezogen und messen in der Regel bestimmte Merkmale oder Eigenschaften nach vorgegebenen, regelhaften Kriterien. Dabei werden Anforderungen der Thesauri zugrunde gelegt. Eine Klassifikation muss keine semantische Zuordnung im Sinne von Vorzugsbezeichnungen oder Relationsbeziehungen auf gleicher Hierarchieebene enthalten. Besonders Taxonomien ordnen Objekte nach bestimmten Rangstufen oder auch metrischen Systemen. Klassifikationen dienen dazu:

- Begriffe einer Entität eindeutig hierarchisch nach einem bestimmten Merkmal zu messen
- Ordnungskriterien und Hierarchien zu entwickeln und
- Analysesysteme für Bezeichnungen und Begriffe für wissenschaftliche Anwendungen und Ordnungssysteme auszugeben

Normdaten/Normdatei (Authority Data): Normdaten sind öffentlich verfügbare, standardisierte kontrollierte Vokabulare, die für die Repräsentation, Beschreibung und Identifikation von bestimmten Entitäten durch eine größere Community gemeinsam genutzt werden und über gemeinsam gepflegte Regeln, persistente Identifikatoren und strukturelle stabile Organisationen betrieben werden. Aufgrund der Bedeutung von Normdaten für die Community sind solche Vokabulare gut dokumentiert und beschrieben, besitzen Möglichkeiten für den Download in verschiedenen Datenformaten und werden durch weitere Serviceleistungen ergänzt.

Ontologie/Schemata: Eine Ontologie ist ein informatisches Konzept, das dazu dient, die Ordnungen von Vokabularen und ihrer Konzepte in ein technisches Schema und formale Modelle zu überführen, um sie über webbasierte Protokolle zu verknüpfen. Dabei beschreiben sie vor allem die Beziehungen zwischen den unterschiedlichen Begriffen und ihren Konzepten bzw. hierarchischen Ordnungsprinzipien.

Sonstiges: Bitte nur als Stichwort.

A5 (1/5) P901: [Anwendbarkeit] Welchem Zweck soll Ihr Vokabular dienen?

Ihre Angabe dient der ungefähren Vorstellung davon, zu welchen Zwecken das Vokabular genutzt werden kann. Das Vokabular dient womöglich allen, beschränken Sie sich in Ihrer Auswahl aber bitte lediglich auf die hauptsächlichen Zwecke.

A6 (1/6) P15: [Medientyp] In welcher Form wurde Ihr Vokabular publiziert?

Ihre Antwort gewährleistet eine Auskunft über die Erreichbarkeit. So ist bei einer Onlinepublikation die Datenübernahme in ein anderes System bzw. eine Weiterverwendung möglich, bei einer Datenbank hingegen nicht.

!A7 (1/7) P64 : [Medientyp] Bitte geben Sie den Literaturbeleg und/oder den onlinebasierten Verweis zu Ihrem Vokabular an.

Gefragt wird nach einer Quellenangabe zu den verschiedenen Details (Zitation) bei Druckpublikation, Datenbank, Textdatei oder ähnlichen Ausgaben.

Vorname Nachname, Titel, Ort Jahr, DOI/Link/URN.

!A8 (1/8) P845: [Urheber] Falls es Autor*innen, Beiträger*innen bzw. Urheber*innen für Ihr Vokabular gibt, führen Sie diese bitte hier auf.

Mit "Urheber" werden hier alle Personen angesprochen, die für die Entstehung des Vokabulars verantwortlich sind und genannt werden sollen. Dabei können gerne unterschiedliche Rollenkonzepte, beispielsweise im Sinne der Contributor Roles Taxonomy [<https://credit.niso.org/>] differenziert werden.

Bitte geben Sie Namen immer in der Form: Doreen von Mühlenberg unter Verzicht auf Titel oder weitere Namenszusätze ein. Die GND-Nummer bzw. die OrcID-Nummer sollen als vollständige Links: <https://d-nb.info/gnd/133366367> statt nur als Nummer 133366367 eingepflegt werden. Die GND-Nummer lässt sich einfach mit einer Namenssuche im Lobid (GND) [<https://lobid.org/gnd>] bzw. im OrcID-System [<https://orcid.org/>] ermitteln. Sollten Sie nicht über eine GND-Nummer verfügen, ist die Ansetzung einer OrcID empfehlenswert. Auch hier bitte den gesamten Link zur OrcID angeben.

A9 1/9 P518: [Technische Realisierung] Wer ist für die technische Umsetzung verantwortlich?

Beispielsweise Personen, die Vokabulare für die Veröffentlichung kuratiert oder vorbereitet haben.

!A10 (1/10) P176: [Herausgegeben von] Von wem wird das Vokabular herausgegeben?

Die Person/Einrichtung, die das Vokabular veröffentlicht hat (z. B. Person, Verlag). Die ISIL-Nummer kann in der Datenbank der Deutschen ISIL-Agentur und Sigelstelle der Staatsbibliothek zu Berlin recherchiert oder beantragt werden [<https://sigel.staatsbibliothek-berlin.de/vergabe/isil>]. Geben Sie hier den Link für die URI ein.

!A11 (1/11) P206: [Verlag] Bitte geben Sie den Verlag oder die Institution an, die das Vokabular bereitstellt.

Die Person/Einrichtung, die das Vokabular technisch bereitstellt und/oder verfügbar hält.

A12 (1/12) P722: [E-mail] Bitte geben Sie eine (funktionale) Mailadresse an, über die eine Kontaktaufnahme möglich ist.

Bei einer funktionalen E-Mail-Adresse handelt es sich um eine Kontaktadresse, die gemeinsam innerhalb einer Funktion/eines Aufgabenbereiches genutzt wird. Diese hat den Vorteil, dass auch bei Abwesenheit einer zuständigen Person oder bei Personalwechsel die Adresse valide bleibt. Das ermöglicht langfristig erfolgreiche Kontaktaufnahmen. Sollte keine funktionale Mailadresse vorhanden sein, geben Sie gerne eine andere Adresse an, unter der man Sie erreichen kann.

!A13 (1/13) P222: [Publikationsdatum] In welchem Jahr wurde Ihr Vokabular erstmalig publiziert? Oder für wann ist die Publikation geplant?

Jahr vierstellig: JJJJ

!A14 (1/14) P608: [Behandeltes Wissensgebiet] Zu welchen (historischen) Wissensfeldern lässt sich Ihr Vokabular zuordnen - Vergeben Sie bitte fünf passende Schlagworte.

Bitte verwenden Sie möglichst präzise Schlagwörter nach eigenem Ermessen, keine weit gefassten Oberbegriffe. Sie sollen eine möglichst gezielte Suche ermöglichen.

!A15 (1/15) P936: [Objekttypen] Welche Entitäten erfasst Ihr Vokabular? (Personen, Orte etc.)

Entität oder Kategorie bezeichnet eine bestimmte Gattung von Begrifflichkeiten, die durch gemeinsame Merkmale definiert werden. So sind bspw. Tätigkeitsprofile und Anforderungsniveaus neben anderen Merkmalen ein Definitionskriterium für die Entität "Berufe". Andere Entitäten sind z. B. Personen, Sachbegriffe, Ortsbezeichnungen, Verwandtschaftsbeziehungen und vieles andere mehr.

!A16 (1/16) P613: [Zahl der integrierten Objekte] Wie viele Objekte (Worte, Begriffe, Normbezeichnungen, Deskriptoren) umfasst Ihr Vokabular?

Erfassen Sie die Anzahl Ihrer normierten Begriffsbezeichnungen sowie der vorhandenen Wort- und Schreibvarianten (bei der Verwendung von Quellen).

!A17 (1/17) P18: [Sprache] In welchen Sprachen liegt Ihr Vokabular vor?

Zur Sprachenauswahl: Siehe Frage A2.

Häufig kommen in Vokabularen auch Übersetzungen weniger einzelner Begriffe vor. An dieser Stelle fragen wir aber lediglich nach einer vollständigen Übersetzung Ihres gesamten Vokabulars oder des größten Teils.

A18 (1/18) P942: [Zahl der Hierarchieebenen] Wie viele Hierarchieebenen weist Ihre Kategorisierung auf?

Hier ist die vertikale Hierarchieebene gemeint, also die hierarchische Untergliederung, beispielsweise vom Allgemeinen ins Spezielle (z. B. bei Dewey Decimal Classification:

200 Religion (1. Hierarchieebene), 270 Geschichte des Christentums (2. Hierarchieebene, gekennzeichnet durch zweite Ziffer)/ es sind die ersten zwei Hierarchieebenen der DDC (Dewey Decimal Classification)⁶.

Zehn Hauptklassen der DDC	Zweite Ebene der Hauptklasse "Religion"
000 Informatik, Information und Wissen, allgemeine Werke	200 Religion
100 Philosophie & Psychologie	210 Religionsphilosophie, Religionstheorie
200 Religion →	220 Bibel
300 Sozialwissenschaften	230 Christentum, christliche Theologie
400 Sprache	240 Christliche Erfahrung, christliches Leben
500 Naturwissenschaften	250 Christliche Pastoraltheologie, Ordensgemeinschaften
600 Technik	260 Kirchenorganisation, Sozialarbeit, Religionsausübung
700 Künste & Freizeit und Erholung	270 Geschichte des Christentums
800 Literatur	280 Christliche Konfessionen
900 Geschichte & Geografie	290 Andere Religionen

[Übersicht DDC: https://www.dnb.de/DE/Professionell/DDC-Deutsch/DDCUebersichten/ddcUebersichten_node.html]

⁶ Dewey Decimal Classification, international am meisten verbreitete Klassifikation zur inhaltlichen Erschließung von Bibliotheksbeständen.

A19 (1/19) P902: [Methodik] Auf welcher Datengrundlage basiert Ihr Vokabular und in welcher Form ist diese nachgewiesen?

Aktuelle Fachbegriffe zeichnen sich dadurch aus, dass sie keine historischen Quellenbegriffe sind, sondern aus der Wissenschaft generierte, verallgemeinerte Fachbegriffe; auch historische Fachbegriffe können also Quellenbegriffe sein, da sie ggf. heute so nicht mehr verwendet werden würden.

Georeferenzierung: Gibt es raumbezogene Informationen zu Ihrem Vokabular? Beispielsweise: Ihr Vokabular zeichnet Denkmäler in Ort X aus, die Denkmäler sind mithilfe von Koordinaten beschrieben, also georeferenziert.

Datierung: Gibt es zeitbezogene Informationen zu Ihrem Vokabular? Beispielsweise: Sie haben ein Vokabular zu Bauwerken in Stadt X. Gibt es eine zeitliche Einordnung zum Bauwerk, z.B. durch das Jahr des Baus/der Einweihung.

Beispiel: Eckard Rolfs Gebrauchstextsorten

- Basiert auf Quellenbegriffen: Ja (Begrifflichkeiten der verschiedenen Gebrauchstextsorten aus historischen Wörterbüchern)
- Basiert auf aktuellen Fachbegriffen: Ja (Nomenklatur und Kategorien der Sprechakttheorie mit Definition)
- Die Quellenbegriffe verfügen über einen Nachweis: Nein (Quellenangabe für Vokabular ist enthalten)
- Die Quellenbegriffe verfügen über eine Datierung: Nein (Die Begriffe besitzen keine Datierung)
- Die Quellenbegriffe verfügen über eine Georeferenzierung: Nein (Es ist keine raumbezogene Information angegeben)
- Die Fachbegriffe verfügen über einen Nachweis: Ja
- Die Fachbegriffe verfügen über eine Datierung: Nein
- Die Fachbegriffe verfügen über eine Georeferenzierung: Nein
- Die Fachbegriffe verfügen über ein theoretisches Konzept: Ja

A20 (1/20) P899: [Eigenschaften des Objekttyps] Welche Eigenschaften der Begriffe werden in Ihrer Kategorisierung abgefragt?

Hier wird das Messkriterium einer Entität abgefragt. Welche Eigenschaften definieren Begriffe oder Klassen? Beim Thema "Berufe" können beispielsweise die Kompetenzen und Skills von Tätigkeiten für Berufsgruppen verglichen werden (z. B.: Brot backen, Mehl mahlen, Teig kneten, Getreide verarbeiten), oder die verschiedenen Qualifikationsstufen des Berufes (z. B.: Gehilfe, Facharbeiter, Geselle, Meister) abgefragt werden.

!A21(1/21) P724: [Beschreibung] Bitte beschreiben Sie Ihr Vokabular kurz.

Mit dieser Beschreibung können Sie Ihr Vokabular präsentieren. Sie soll die Nutzer*innen in die Lage versetzen, schnell erfassen zu können, ob das Vokabular für sie von Interesse sein kann.

Bitte schreiben Sie allgemein verständlich, damit auch Forschende außerhalb Ihrer Fachwissenschaft davon profitieren.

Zur Orientierung, welche Elemente die Beschreibung erfassen sollte:

- Genauere Beschreibung der eigentlichen Thematik des Vokabulars und seiner üblichen Gebrauchszwecke. Dabei können auch abgefragte Details zum Umfang, zur Entität, zur Anzahl der Objekte, Vorgänger oder Cross-Konkordanzen, Urheber und institutionelle Anbieter etc. näher beschrieben werden.
- Die Beschreibung soll maximal 1.500 Zeichen (mit Leerzeichen) umfassen

Beispiel 1: Vokabular zu "Die Funktionen der Gebrauchstextsorten" von Eckard Rolf:

Das Vokabular (Stand August 2023) umfasst etwas über 2.100 Gattungsbegriffe zu Gebrauchstexten von der „Mahnung“ bis zum „Entlassungsschein“. Der Datensatz erschien analog in der Publikation von Eckard Rolf, Die Funktionen der Gebrauchstextsorten, Berlin/New York 1993; und wird in digitaler Form in FactGrid (RDF-Datenbank, Wikibase) angeboten. Unter Sprechakt-pragmatischen Gesichtspunkten werden in fünf zunehmend nuanciert ausdifferenzierten Hierarchien nicht-literarische Textsorten klassifiziert. Das Vokabular eignet sich für quantitative und qualitative Analysen größerer Textkorpora.

Beispiel 2: Ontologie der historischen, deutschsprachigen Amts- und Berufsbezeichnungen:

Das taxonomische, kontrollierte Vokabular systematisiert Tätigkeiten verschiedener Berufe oder äquivalenter Bezeichnungen (Tätigkeiten, Firmenbezeichnungen) hinsichtlich ihrer Berufsfachlichkeit, Skills und Kompetenzen und ordnet sie in sechs verschiedene Hierarchien (Normbezeichnung, fünf Hierarchieebenen) ein. Zudem werden berufsbezogene Anforderungsniveaus messbar (keine sozialökonomischen Hierarchien). Das Vokabular stellt eine deutschsprachige, nationale Ableitung vom internationalen Standard der ISCO 08 ab und wurde von der Agentur für Arbeit als "Klassifikation der Berufe 2010/2020" (KldB 2010) grundlegend erarbeitet und methodisch entwickelt. Während die KldB jeweils eine dynamische Anpassung an die sich veränderten Berufslandschaften vornimmt, wird mit OhdAB ein stabiles, dauerhaftes Ordnungssystem für historische Berufs- und Amtstätigkeiten angestrebt. Es geht vom "Berufsstand" der historischen Massenregister (Kirchenbücher, Adressverzeichnisse, Steuerregister etc.) aus und erfasst daher auch Angaben (Familienbeziehungen, Bezeichnungen für Vaganten, Mitglieder in Organisationen etc.), die heute nicht als Berufsangaben gelten. Anders als die KldB 2010/20 werden dazu die Varianten der einzelnen Berufsbezeichnungen transparent zu den Normbezeichnungen zugeordnet. Dies ermöglicht ein teilau-

tomatisiertes Matching. Das Vokabular ermöglicht die Detektion weiblicher und männlicher Berufsbezeichnungen und besitzt einen größeren Anteil englischsprachiger Übersetzungen.

A22 (1/22) P724engl: [Beschreibung Englisch] Bitte geben Sie hier, sofern vorhanden, die Beschreibung Ihres Vokabulars auf Englisch an. Siehe Frage 23

Falls Sie keine eigene Übersetzung anfertigen möchten, sind Übersetzungsprogramme wie [DeepL](#) oder [Chat.GPT](#) hierfür gut nutzbar.

A23 (1/23) P887: [Datensatzlink] Bitte geben Sie hier den Link zur Beschreibung Ihres Vokabulars an, wenn vorhanden.

Vorteilhaft ist eine persistente Veröffentlichung mit DOI bspw. über Zenodo oder andere Publikationsplattformen, ansonsten sind hier Beschreibungen auf Projektwebsites gemeint.

A24 (1/24) P243: [Zentraler Gegenstand] Wie würden Sie Ihr Vokabular mit einem Satz zusammenfassen?

Beispiel 1: Über 2100 Gattungsbegriffe zu nicht-literarischen Gebrauchstexten und ihren Funktionen nach Eckard Rolf (Stand August 2023).

Beispiel 2: Klassifikation nach Tätigkeitsprofilen und Anforderungsniveaus auf Grundlage der KldB 2010 für historische Berufe.

!A25 (1/25) P111: [Adresse als Zeichenfolge] Bitte geben Sie hier die Adresse der herausgebenden Einrichtung an, sofern vorhanden.

Teil B: Zugänglichkeit

B26 (2/1) P125: [Zugänglichkeit] Wie ist Ihr Vokabular zugänglich?

Die Frage dient dem Ziel, Zugänglichkeit und Einschränkungen der Benutzbarkeit zu erfassen. Vermerkt werden auch kostenpflichtige Angebote.

!B27 (2/2) P180 [Lizenzierung] Bitte geben Sie hier die Lizenzierung Ihres Vokabulars an.

Erläuterungen zu den Lizenzen finden sich auf

- <https://creativecommons.org/share-your-work/ccllicenses/>
- <https://rightsstatements.org/page/1.0/?language=de>

B28 (2/3) P939: [Versionierung] Verfügt Ihr Vokabular über eine Versionierung?

Werden Bearbeitungsstände durch einzelne Fassungen kenntlich gemacht und unterschieden?

B29(2/4) P818: [Erneuerungsdatum] Wann wurde Ihr Vokabular das letzte Mal aktualisiert?

Bei Jahr bitte: JJJJ; bei monatsgenauer Angabe: MM.JJJJ; bei tagesgenauer Angabe: TT.MM.JJJJ.

B30 (2/5) Bed001: [Dokumentation] Existiert eine Dokumentation Ihres Vokabulars?

Unter Dokumentation wird eine umfassende Beschreibung des Vokabulars mit Angaben zu Entstehung, Methodik, Arbeitsweisen, Quellengrundlagen etc. verstanden.

B31 (2/6) P382 [Spezifiziert in Frage 30] (Bedingung Fr. 30 Ja):

Bitte geben Sie im Sinne einer Zitation die Daten der Urheber*innen der Dokumentation ein und folgen Sie dabei dem Muster: Vorname Nachname, Titel, Ort Jahr, DOI/Link/URN. Mit Urheber werden hier alle Personen angesprochen, die für die Entstehung des Vokabulars verantwortlich sind und genannt werden sollen. Dabei können gerne unterschiedliche Rollenkonzepte, beispielsweise im Sinne der Contributor Roles Taxonomy [<https://credit.niso.org/>] differenziert werden.

B32 (2/7) Bed002: [Tutorial] Verfügt Ihr Vokabular über die Dokumentation hinaus über ein Tutorial oder eine Gebrauchsanweisung?

Damit sind alle Hilfsmittel gemeint, die ein tieferes Verständnis zum Aufbau, zu den technischen Grundlagen, zur Verwendung oder zur weiteren Ausgestaltung geben können und auf Lehr- und Lernprozesse ausgerichtet sind.

B33(2/8) P978 (Bedingung Frage 32 Ja oder Teilweise) [Externes Tutorial] Bitte geben Sie die Links oder vollständigen Zitationen zu diesen Lernmaterialien an.

Bitte geben Sie im Sinne einer Zitation die Daten der Materialien ein und folgen Sie dabei dem Muster: Vorname Nachname, Titel, Ort Jahr, DOI/Link/URN.

B34 (2/9) P900: [Vollständigkeit] Ist Ihr Vokabular vollständig oder auf Vollständigkeit angelegt?

Ermittelt werden soll, wie stark die Konzeption eines Vokabulars auf eine vollständige Erfassung und Beschreibung einer Entität ausgerichtet ist. Dies soll den Nutzenden einen Eindruck zur Vollständigkeit des Vokabulars und zu etwaigen Weiterverwendungsmöglichkeiten geben.

Faktisch vollständig: das Vokabular ist abgeschlossen (Bsp.: Das Vokabular beschreibt alle Kirchen der Stadt X im Jahre Y; alle Kirchen sind bereits eingetragen, es ist nicht davon auszugehen, dass das Vokabular noch erheblich erweitert werden muss/wird.

Vollständigkeit wird angestrebt: das Vokabular ist noch nicht abgeschlossen, aber es ist ein realistisches Ziel, Vollständigkeit zu erreichen. (Bsp.: Das Vokabular soll alle Kirchen der Stadt X im Jahre Y beschreiben; einige sind bereits enthalten, andere werden noch beschrieben und ergänzt werden).

Vokabular ist für Wachstum offen: Vokabular ist erweiterbar (Bsp.: Denkmäler der Region XY; bei Entdeckung/Errichtung neuer Denkmäler kann das Vokabular entsprechend erweitert werden).

Bildet einen Teil ab: Das Vokabular strebt keine Vollständigkeit an.

B35 (2/10) P999: [Pflege] Wie wird Ihr Vokabular gepflegt?

B36 (2/11) P1000: [Änderungen] Von wem können Änderungen an Ihrem Vokabular vorgenommen werden?

B37 (2/12) P575: [Aspekte der Gestaltung] Nach welchen Prinzipien erfolgt die begriffliche Erschließung?

Regelwerke definieren, nach welchen Regeln die jeweiligen Vokabulare erstellt und beschrieben sind, einerseits in Bezug auf den Inhalt des Vokabulars (Was wird aufgenommen ...), andererseits auf die Form des Vokabulars (z. B. Schreibweisen) oder das Datenformat. Ein Regelwerk dient der Standardisierung, Normierung und damit der Verbesserung der Konsistenz der Vokabulare. Das erhöht die Auswertbarkeit, Qualität und hilft Dubletten zu vermeiden. Außerdem ist die Standardisierung des Datenformats relevant für eine einheitliche Repräsentation sowie für möglichst barrierefreie Datenübernahmen oder den Datenaustausch. Arbeitsumgebungen von Gedächtniseinrichtungen definieren Regelwerke als Hilfsmittel der Katalogisierung (z. B. RSWK) und Erfassung, die Nutzeranforderungen spezifizieren (FRBR, FRISAD) oder allgemeine Indexierungsrichtlinien festschreiben (DIN-Normen) bzw. Standardisierungsrichtlinien formulieren (z. B. ISO 25964, SKOS).⁷

Teil C: Crosskonkordanzen

C38 (3/1) P898: [Crosskonkordanz] Mit welchen anderen Vokabularen (GND-Entitäten, Wikidata-Objekten etc.) ist ihr Vokabular gematcht? Wenn es Ihnen möglich ist, geben Sie bitte eine Schätzung ab, wieviele Ihrer Daten konnten Sie im gematchten Datenbestand auffinden und wieviele Daten aus dem gematchten Datenbestand sind in Ihren Daten enthalten?

Beispiel:

Die Ontologie historischer Berufsbezeichnungen wurde mit der GND abgeglichen, also gematcht. Alle Berufsbezeichnungen der GND sind in der Ontologie enthalten, aber nur ca. 9% der in der Ontologie enthaltenen Begriffe finden sich in der GND.

⁷ Jutta Lindenthal: Regelwerke, in: DigiCult.xTree (Hg.): GND-KIT,, 2024, URL: <https://digicult.atlas-sian.net/wiki/spaces/GNDKIT/pages/584122713/Regelwerke>.

	Name Zielvokabular	Matching der Begriffe aus dem Zielvokabular mit Ihrem eigenen (Vollständig, Teilweise, Geringfügig)	Anzahl der Begriffe aus dem Zielvokabular in Ihrem eigenen (Geschätzte Vollständigkeit in %)	Matching der Begriffe aus Ihrem eigenen mit Zielvokabular (Vollständig, Teilweise, Geringfügig)	Anzahl der Begriffe aus Ihrem eigenen im Zielvokabular vorhanden (Geschätzte Vollständigkeit in %)	Sprache des abgeglichenen Vokabulars
1.	GND	Vollständig	100	Geringfügig	9	de
2.	Klassifikation der Berufe 2010	Vollständig	100	Teilweise	60	de
3.	Classification of Occupations (ISCO)	Vollständig	100	Geringfügig	7	de
4.	Classification of Occupations (ISCO)	Geringfügig	20	Vollständig	100	en
5.						
6.						
7.						
8.						

C39 (3/2) P937: [Identifizier] Verfüg Ihr Vokabular über eindeutige und stabile Identifizier?

Identifizier oder Identifikatoren dienen zur eindeutigen Identifizierung;

Interne Identifizier

Identifikatoren, auf die lediglich intern zugegriffen werden kann z. B. in Ihrer Datenbank oder Ihrer Printpublikation zu finden sind.

Persistente Identifizier

Identifikatoren, die im Web vorliegen und über die ein dauerhafter Zugriff gewährleistet ist.

Auf URLs basierende Identifizier

Identifikatoren, die im Web vorliegen und über URL lokalisiert werden, aber auf die kein dauerhafter Zugriff gewährleistet ist (zieht die Webseite beispielsweise um, erhält man einen "toten" Link und der Zugriff ist nicht mehr möglich).

Auf URIs basierende Identifizier

Persistenter, dauerhafter Zugriff gewährleistet, auf Grundlage von URI.

Klassifizierende Identifizier

Identifikatoren, die in Klassen sortiert sind und damit eine Klassifikation bilden.

C40 (3/3) Bed003: [Nutzendengruppen Ja/Nein] Können Sie Auskunft über Ihre Nutzendengruppe geben?

C41 (3/4) P1095 [Nutzendengruppen] (Bedingung Fr. 40 Ja): Falls es Ihnen bekannt ist, geben Sie hier bitte an, von wem Ihr Vokabular extern bereits genutzt wurde.

Teil D: Bereitstellung der Daten

D42 (4/1) P940: [Technologie] Mit welcher Technologie/Software wird Ihr Vokabular verfügbar gemacht (RDF-Datenbank, Skosmos, Dante, xTree,...)?

!D43 (4/2) P903: [Datenformat] Bitte geben Sie hier das Datenformat/die Datenformate und Exportmöglichkeit Ihres Vokabulars an. (CSV, Excel, PDF, Skos, RDF, Json, xml, ...)

Verwenden Sie bitte die üblichen Dateiendungskürzel der Dateiformate (drei- oder vierstellig).

!D44 (4/3) P923: [API] Wenn vorhanden, geben Sie hier bitte die API-Schnittstelle Ihres Vokabulars als URL an.

APIs verwenden häufig bestimmte Datenformate für Anfragen und Antworten, wie JSON (JavaScript Object Notation) oder XML (Extensible Markup Language). Das Untersuchen der Struktur der Anfragen und Antworten kann darauf hinweisen, dass es sich um eine API handelt. Bitte geben Sie die API auch an, wenn die Nutzung eine Authentifizierung erfordert. Solche Angaben finden Sie auch in den Dokumentationen und technischen Berichten Ihres Angebotes.

!D45 (4/4) P941: [SPARQL-Endpunkt URL] Bitte geben Sie hier, sofern vorhanden, den SPARQL-Endpunkt Ihres Vokabulars an.

Dieser ermöglicht eine Abfrage der Daten mit SPARQL. Ein SPARQL-Endpunkt hat normalerweise eine spezifische URL, an die SPARQL-Abfragen gesendet werden können. Diese URL endet oft mit "sparql" oder "query". SPARQL-Endpunkte werden normalerweise mit einer Dokumentation bereitgestellt, die erklärt, wie man SPARQL-Abfragen sendet, welche Abfragesprache unterstützt wird und welche Daten verfügbar sind.

D46 (4/5) P995: [Service-Angebote] Welche Services werden zu Ihrem Vokabular angeboten? (Zum Beispiel Beratungsservice, Matchingservice, Nachcodierung etc...)

Unter Services können sowohl inhaltliche wie technische Angebote verstanden werden. So wären beispielsweise Beratungsservice, Matchingservice oder Angebote zur Erweiterung einer Klassifikation denkbar.

Teil E: Ausblick

E47(5/1) P263: [Arbeitsstand] Hier können Sie Informationen zum Bearbeitungsstand Ihres Vokabulars angeben (z. B. in Arbeit, einzelne Sprachen noch nicht vollständig übersetzt etc.).

E48 (5/2) P73: [Evaluationsfrage] Werden in unserem Fragebogen Informationen nicht erfasst, die Sie für wichtig halten und die noch ergänzt werden sollten?