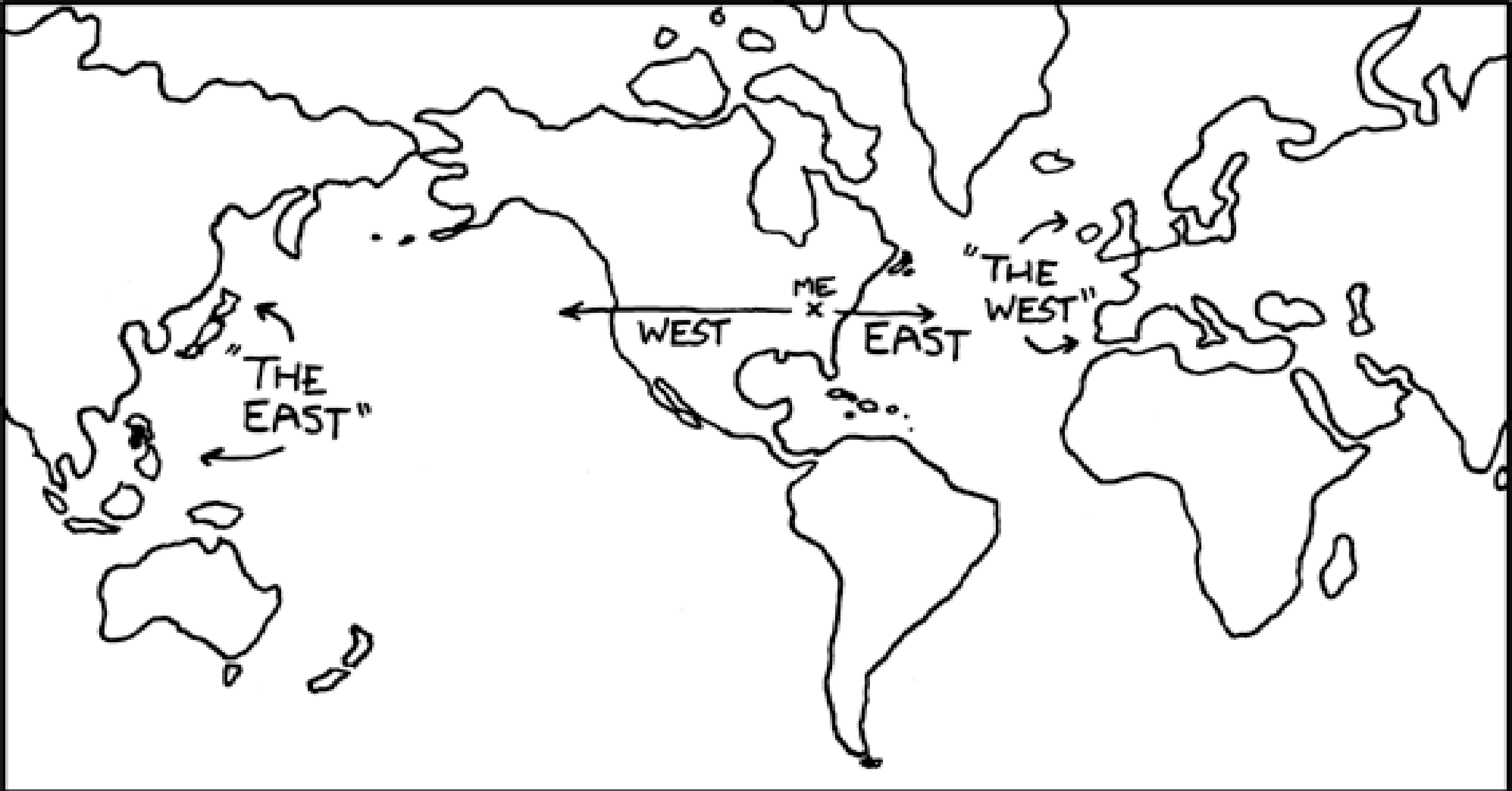




Forschungsdaten Struktur geben: eine Einführung in Ontologien, Terminologien und kontrollierte Vokabulare

Jessica Rex

08.05.2023



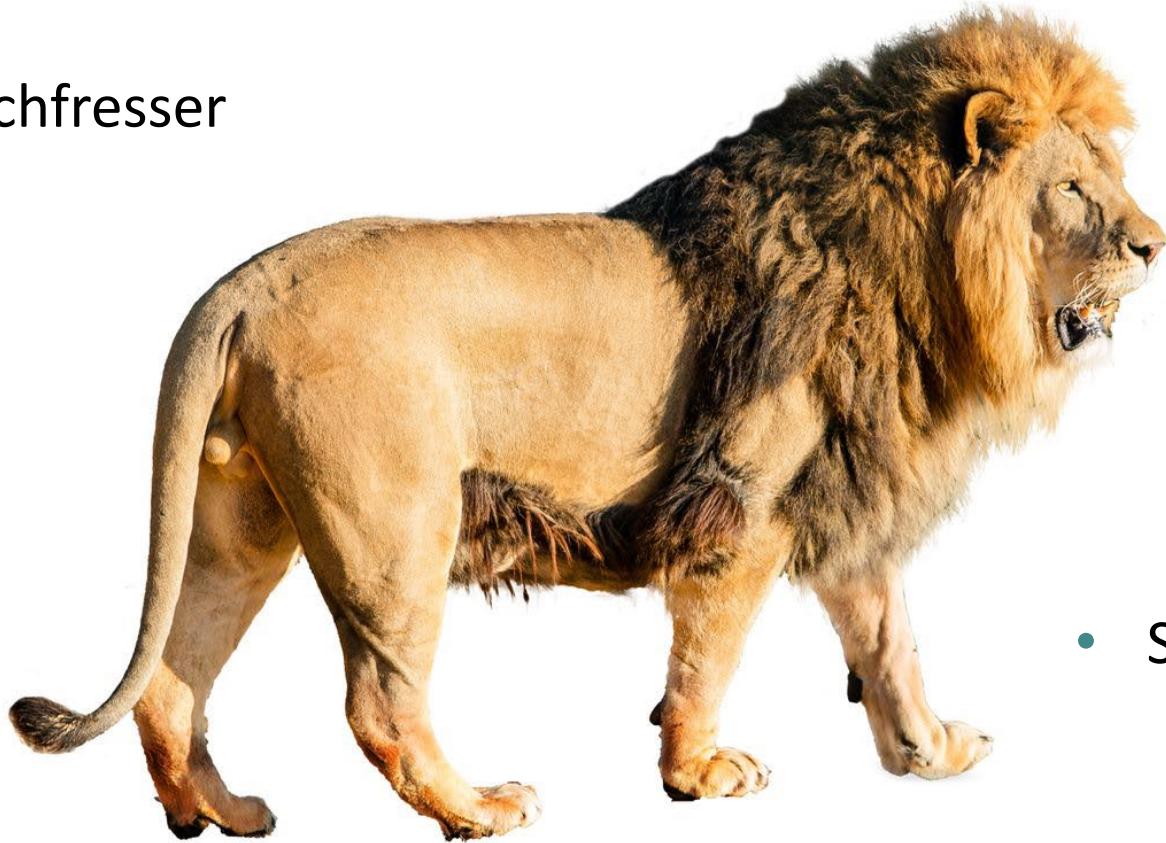
THIS ALWAYS BUGGED ME.

Wie würden Sie dieses Objekt beschreiben?

- Fleischfresser

- Mähne

- Fell



- Großkatze

- Savannenbewohner

- Säugetier

Begriffe: Terminologie

- „Gesamtheit aller Fachwörter (Termini) eines Fachgebietes“
- möglichst exakte Definitionen, Synonyme und ergänzende Informationen
 - hierarchische Organisation, Baumstruktur (= bspw. **Thesaurus**, Klassifikation)
 - Abbildung von Domänenwissen durch Konzepte und Relationen in einer Netzstruktur (= bspw. **Ontologie**)
- enthalten **kontrolliertes Fachvokabular**
- können natürlich-sprachlich oder semantisch-formell sein

Wozu brauchen wir Terminologien in der Forschung?

1. Sie befördern eine effiziente und eindeutige fachsprachliche Kommunikation.
2. Sie sind Grundlage für den Wissenstransfer.
3. Sie sind unabdingbar für die sprachliche Qualitätssicherung in der Forschung.
4. Sie verbessern die Wiederauffindbarkeit von Forschungsdaten (durch semantische Indizierung).
5. Sie verbessern die Nachnutzbarkeit von Forschungsdaten (durch angereicherte Dokumentation).

Begriffe: Kontrolliertes Vokabular

- natürlich-sprachliche, standardisierte und organisierte Anordnungen von Wörtern und Phrasen
- inhaltliche Beschreibung von Objekten wie bspw. Forschungsdaten
- Verwendung in Metadatenschemata
- strukturierte Dokumentation von Daten

Schlagwortliste

- Tier
- Säugetier
- Omnivore
- Canidae
- ...

Taxonomie

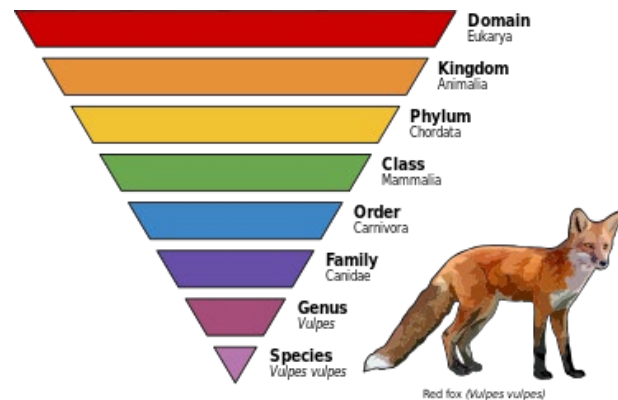


Image by https://commons.wikimedia.org/wiki/User_talk:Nninnabee

Thesaurus

THESAURUS English Web 2015 (enTenTen15)

frustrating as adjective 94,305x

Word	Frequency ?		
1 confusing	118.003	6 scary	157.063
2 stressful	96.758	7 daunting	82.616
3 tedious	50.596	8 embarrassing	58.848
4 frightening	65.109	9 disturbing	113.509
5 disappointing	103.245	10 painful	248.801

Rows per page: 10 1-10 of 100 1

Fachspezifische Thesauri

- Agrarwissenschaften: AGROVOC Multilingual agricultural thesaurus, <http://aims.fao.org/vestregistry/vocabularies/agrovoc-multilingual-agricultural-thesaurus>
- Geisteswissenschaften: A Thesaurus of Old English, <http://oldenglishtesaurus.arts.gla.ac.uk/>
- Informationswissenschaften: INFODATA Thesaurus, http://www.infodata-edepot.de/thesaurus/T_SM.HTM
- Kunst und Architektur: Art and Architecture Thesaurus (AAT), <http://www.aat-deutsch.de/>
- Lebenswissenschaften: Umweltthesaurus (UMTHES), <https://sns.uba.de/umthes/de.html>
- Medizin und Biowissenschaften: Thesaurus Medical Subject Headings (MeSH), <https://www.nlm.nih.gov/mesh/>
- Psychologie: Thesaurus Psychologie (PSYINDEX), <http://www.zpid.de/index.php?wahl=products&uwahl=printed&uuwahl=psyndexterms>
- Wirtschaftswissenschaften: Standard Thesaurus Wirtschaft (STW), <http://zbw.eu/stw/version/latest/about>

Beispiele für Vokabularien

Geisteswissenschaften

- Rare Books and Manuscripts Section (RBMS) Controlled Vocabularies
- Union List of Artist Names (ULAN)

Natur- und Lebenswissenschaften

- International Classification of Diseases (ICD)

Sozial- und Verhaltenswissenschaften

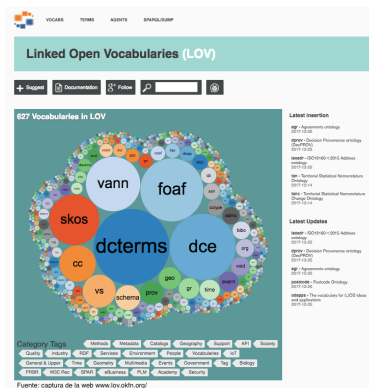
- Data Documentation Initiative (DDI) Controlled Vocabularies

BARTOC

- mehr Klassifikationen, Thesauri und kontrollierte Vokabulare finden Sie im **Basel Registry of Thesauri, Ontologies & Classifications** (<https://bartoc.org>)



- und bei **Linked Open Vocabularies** (<https://lov.linkeddata.es/dataset/lov/>)

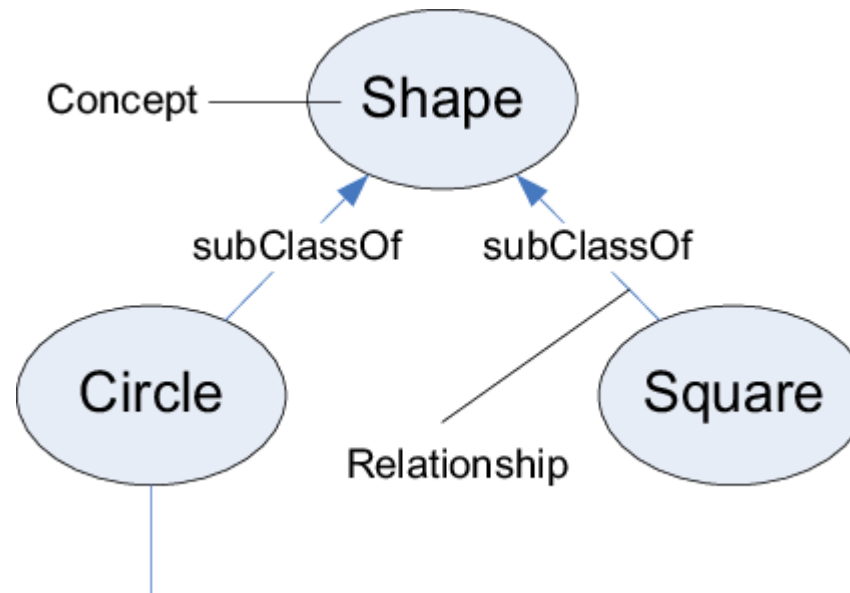


Thesaurus.

Image by Ben Illis <https://www.benillis.com/>

Begriffe: Ontologie

- griechisch „die Lehre des Seins“
- Semantisch-formale Begriffssystematik
- geordnete Menge von Fachbegriffen und deren Beziehungen zueinander
- besteht aus **classes/concepts** (Fachtermini) und **relationships/properties** (die formalisierte Beziehung in welcher diese Termini zueinander stehen) sowie **instances/individuals**



Wie funktioniert eine Ontologie?

- Remember this guy?



Wie funktioniert eine Ontologie?

Ontology: AnimalOntology

Classes:

- Animal
- Mammal
- Bird
- Reptile
- Characteristic
- Behavior

Properties:

- hasHabitat
- hasDiet
- hasCharacteristic
- hasBehavior

Instances:

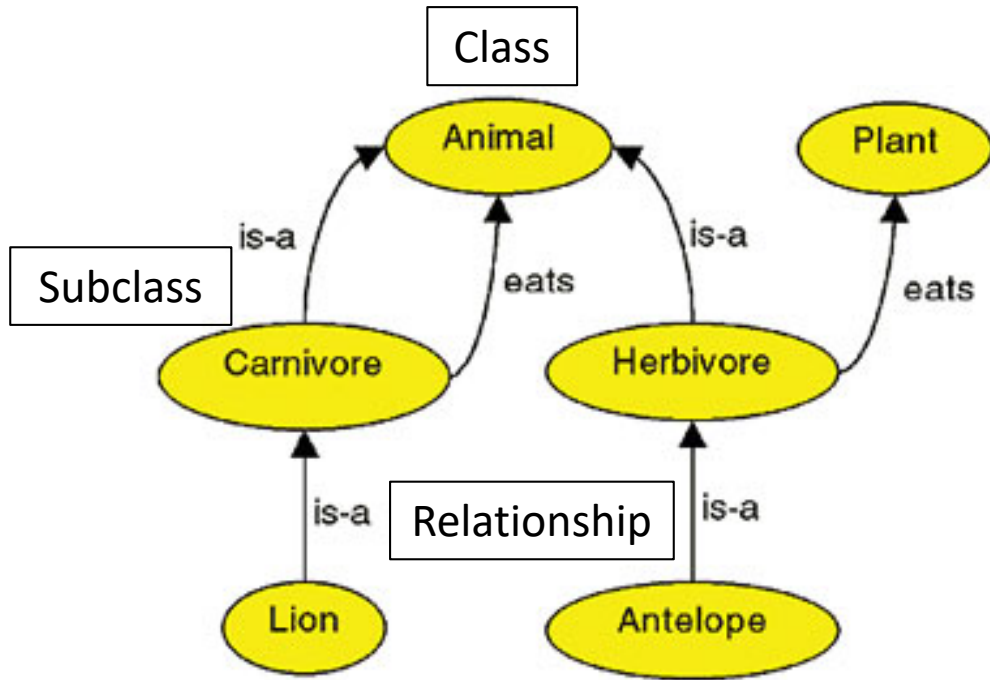
- Animal instances (Lion, Elephant, Penguin, Snake, etc.)
- Habitat instances (Savanna, Forest, Ocean, Desert, etc.)
- Diet Type instances (Herbivore, Carnivore, Omnivore, etc.)
- Characteristic instances (Fur, Feathers, Scales, etc.)
- Behavior instances (Nocturnal, Diurnal, Migratory, etc.)

Relationships

Each Animal instance has a specific habitat (hasHabitat property).
Each Animal instance has a specific diet type (hasDiet property).



Wie funktioniert eine Ontologie?



Instance



Ontologien finden

- im TIB Terminology Service (in Kooperation mit der NFDI) (<https://terminology.tib.eu/ts>)
- Im Archivio Ontology Archive (OWL Ontologien) (<https://archivio.dbpedia.org/list>)

List of all available ontologies

Download Triples Archivio Star Rating Crawling-Status Switch: [Published Web Versions](#) | [Developer Versions](#)

Additional columns: Databus URI Source Semantic Version Parsing Min.License Good License Consistency LODe conformity Latest Timestamp Addition Date

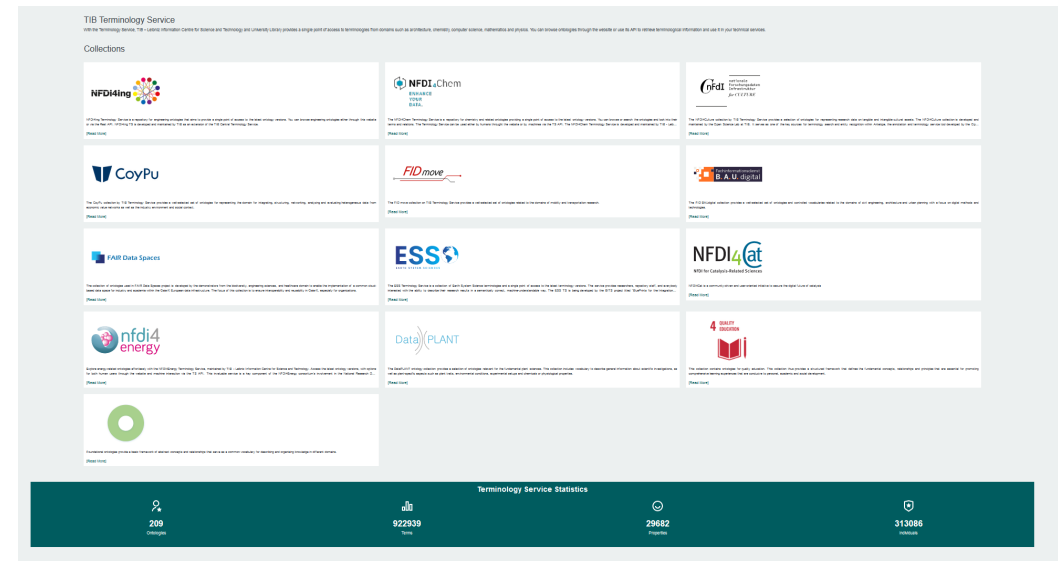
Search

View Archived Ontology	Download Latest	Triples	Stars	Crawling Status
Cinelab ontology	owl, ttl, nt	329	☆☆☆☆	✓
The DBpedia Archivio Test Ontology	owl, ttl, nt	8	☆☆☆☆	✓
http://archivi.ibr.regione.emilia-romagna.it/ontology/eac-cpf/eac-cpf.rdf	owl, ttl, nt	269	☆☆☆☆	✓
http://assemblee-virtuelle.github.io/grands-voisins-v2/gv.owl.ttl	owl, ttl, nt	285	☆☆☆☆	✓
http://assemblee-virtuelle.github.io/grands-voisins-v2/thesaurus.ttl	owl, ttl, nt	101	☆☆☆☆	✓
BabelNet	owl, ttl, nt	10	☆☆☆☆	✗
http://bag.basisregistraties.overheid.nl/def/bag	owl, ttl, nt	1993	☆☆☆☆	✗
Atom Syndication Ontology	owl, ttl, nt	392	☆☆☆☆	✗
The 3D Modeling Ontology	owl, ttl, nt	117	☆☆☆☆	✓
ExtruOnt	owl, ttl, nt	1071	☆☆☆☆	✓
http://biohackathon.org/resource/falco	owl, ttl, nt	235	☆☆☆☆	✓
Basisregistratie Kadaster vocabulaire	owl, ttl, nt	165	☆☆☆☆	✗
top10nl vocabulaire	owl, ttl, nt	7242	☆☆☆☆	✗

Forschungsdaten Struktur geben: eine Einführung in Ontologien, Terminologien und kontrollierte Vokabulare

Der TIB Terminology Service

- Repository für Terminologien, kontrollierte Vokabulare oder Ontologien
- zentraler Einstiegspunkt für disziplin- oder projektspezifische Kollektionen
- durchsuchbare Benutzeroberfläche
- Zugriff per API möglich
- wird in der nächsten Zeit um Entwicklungs- und Kuratierungswerkzeuge, Visualisierungswerkzeuge, Archivierungsservices, Benachrichtigungsdienste und Veröffentlichungsdienste ergänzt



Weiterführende Materialien

- [Metadaten und Metadatenstandards | forschungsdaten.info](https://www.forschungsdaten.info)
- Axel Merced, The role of ontologies in data management, <https://www.linkedin.com/pulse/role-ontologies-data-management-alex-merced-nfile/>
- Train-the-trainer Konzept zum Thema Forschungsdatenmanagement, Einheit 8 Dokumentation und Metadaten: <https://doi.org/10.5281/zenodo.1215377>
- Oliver Koepler et.al., Terminologien, Metadaten, Forschungsdaten - Dienste für Forschungsdaten-Infrastrukturen, urn:nbn:de:0290-opus4-179165

Nächste Coffee Lectures

Das Thüringer Kompetenznetzwerk
Forschungsdatenmanagement präsentiert...

Forschungsdatenmanagement
Online Coffee Lectures 2024



08.05.2024	Forschungsdaten Struktur geben: eine Einführung in Ontologien, Terminologien und kontrollierte Vokabularien Jessica Rex (TU Ilmenau)
29.05.2024	Forschungsdaten durch Ontologien und Standards effizient beschreiben und nachhaltig nutzen – Ein Beispiel aus den Erdsystemwissenschaften Angelina Kraft (TIB Hannover)
26.06.2024	Herausforderungen bei der Arbeit mit Ontologien und deren Anwendung im Forschungsdatenmanagement Felix Engel (TIB Hannover)
24.07.2024	Enabling Semantic Search and Interoperability within NFDI4Biodiversity using BiodivPortal Dr. Naouel Karam (InfAI e.V., Leipzig)
31.07.2024	Ein um Metadaten angereichertes File System (MEF) Carsten Schäuble (Zuse Institute Berlin)



Mehr Informationen und die Anmeldung finden Sie auf:
www.forschungsdaten-thueringen.de

- 29.05. Forschungsdaten durch Ontologien und Standards effizient beschreiben und nachhaltig nutzen – ein Beispiel aus den Erdsystemwissenschaften
- 26.06. Herausforderungen bei der Arbeit mit Ontologien und deren Anwendung im Forschungsdatenmanagement
- 24.07. Enabling Semantic Search and Interoperability with NFDI4Biodiversity using BiodivPortal
- 31.07. Ein um Metadaten angereichertes Filesystem (MEF)

Vielen Dank für Ihre Aufmerksamkeit.