



20 Jahre Semantic Web – eine vergessene oder unnütze Technologie?

Content

Language

Data Management



Acolada Marken



UniLex

Mehrsprachige
Wörterbücher



My fabulous agent

Adobe Indesign Plug-
In für die
automatische
Verarbeitung von
XML in InDesign



SimQin Editor

Einfach zu
verwendender
Wordähnlicher XML-
Editor



Sirius CMS

XML/SGML-Content
Management

Asset-Management und
Dokument-Management



UniTerm

Terminologie
Management



Velingua

Bausteine für
Terminologie



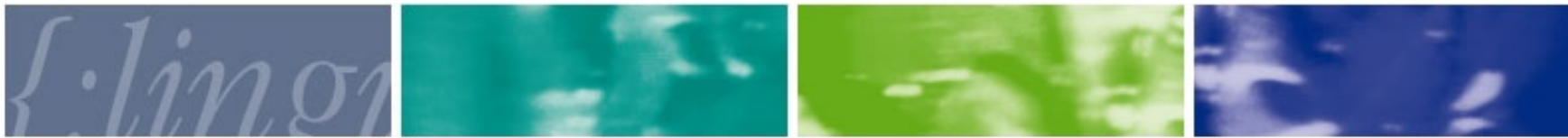
Cobriili

OnlineHilfe
eLearning
Content Delivery Plattform



Was bedeutet Semantic-Web

- > 2001 wurde der RDF-Standard veröffentlicht
(Resource Description Framework)
- > Die Basis des Systems ist das Triple
 - > Subjekt Prädikat Objekt
 - > München ist eine Stadt
 - > München liegt in Bayern
- > Es gibt viele zugehörige Standards, das macht es unübersichtlich am Anfang
- > viel XML, viel Namespaces <ac:style>
- > aber auch Formulierungen von Triples ohne XML sind verfügbar/möglich
(Turtle, JSON)



Hilfreiche Links zu Standards

<https://www.w3.org/RDF/> -> Basis

<https://www.w3.org/TR/rdf-schema/> -> RDFS -> RDF-Schema

<https://www.w3.org/OWL/> -> OWL -> Web Ontology Language

<https://www.w3.org/2004/02/skos/> -> SKOS ->

Simple Knowledge Organization System

Aber auch DCTerms, FOAF, VCard, VANN, ... alle mit dem Bestreben möglichst viele generell abzudecken, damit man sich daran bedienen kann



Tools zum Einstieg

> Protegé -> liest RDF's und kann die anzeigen frei verfügbar -> Beispiel rdf

The screenshot shows the Protegé application window titled "ISO3166-2-SubdivisionCodes-BG". The interface includes a menu bar (File, Edit, View, Reasoner, Tools, Refactor, Window, Ontop, Mastro, Help), a search bar, and several tabs for "Active ontology", "Entities", "Individuals by class", "DL Query", and "Individual Hierarchy Tab".

The main area is divided into two panes:

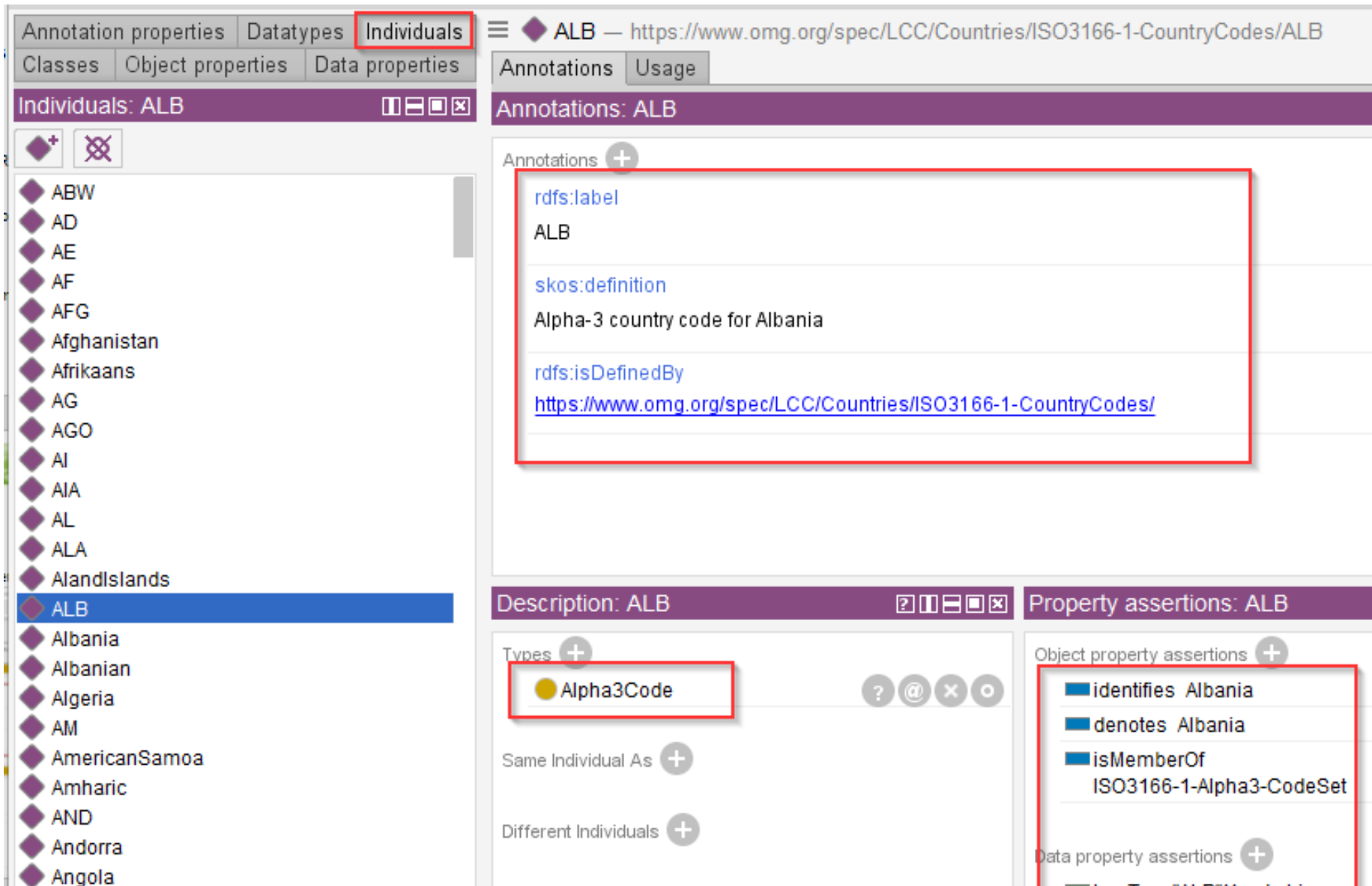
- Left Pane (Class Hierarchy):** Shows a tree structure of classes under "owl:Thing". The "BibliographicCode" class is highlighted in blue. Its parent is "LanguageIdentifier", which is a subclass of "CodeElement".
- Right Pane (Annotations and Description):** Shows the "Annotations: BibliographicCode" section with a red box highlighting the following annotations:
 - rdfs:label:** bibliographic code
 - skos:definition:** [type: xsd:string] an alpha-3 code that is a member of the set of bibliographic codes in ISO 639-2
 - rdfs:isDefinedBy:** <https://www.omg.org/spec/LCC/Languages/LanguageRepresentation/>

Below the annotations is the "Description: BibliographicCode" section, which includes:

- Equivalent To:** (empty)
- SubClass Of:** A red box highlights the following subclasses:
 - Alpha3Code
 - LanguageIdentifier
- General class axioms:** (empty)
- SubClass Of (Anonymous Ancestor):**
 - denotes **only** Language
 - identifies **only** Language



Tools zum Einstieg



Annotation properties | Datatypes | **Individuals** | ALB — <https://www.omg.org/spec/LCC/Countries/ISO3166-1-CountryCodes/ALB>

Classes | Object properties | Data properties | Annotations | Usage

Individuals: ALB | **Annotations: ALB**

Annotations +

- rdfs:label
ALB
- skos:definition
Alpha-3 country code for Albania
- rdfs:isDefinedBy
<https://www.omg.org/spec/LCC/Countries/ISO3166-1-CountryCodes/>

Description: ALB | **Property assertions: ALB**

Types +

- Alpha3Code

Same Individual As +

Different Individuals +

Object property assertions +

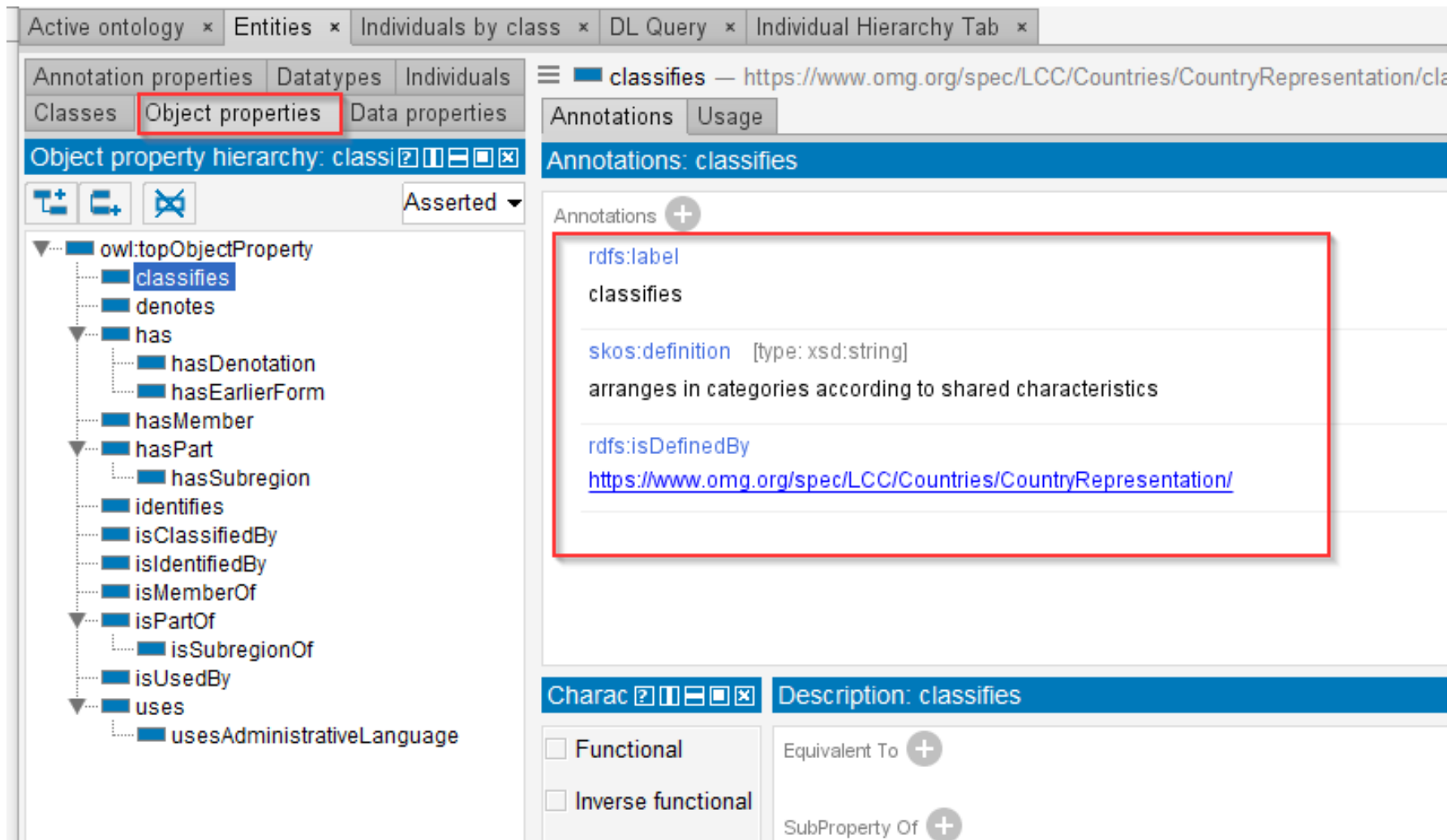
- identifies Albania
- denotes Albania
- isMemberOf
ISO3166-1-Alpha3-CodeSet

Data property assertions +



Tools zum Einstieg

> Protege -> liest RDF's und kann die anzeigen



The screenshot shows the Protege ontology editor interface. The top navigation bar includes tabs for 'Active ontology', 'Entities', 'Individuals by class', 'DL Query', and 'Individual Hierarchy Tab'. Below this, there are tabs for 'Annotation properties', 'Datatypes', and 'Individuals'. The 'Object properties' tab is selected and highlighted with a red box. The main area displays the 'Object property hierarchy: classifies' and the 'Annotations: classifies' section. A red box highlights the following annotations:

- rdfs:label**
classifies
- skos:definition** [type: xsd:string]
arranges in categories according to shared characteristics
- rdfs:isDefinedBy**
<https://www.omg.org/spec/LCC/Countries/CountryRepresentation/>

At the bottom, the 'Description: classifies' section is visible, showing options for 'Functional' and 'Inverse functional' (both unchecked), and 'Equivalent To' and 'SubProperty Of' (both with plus signs).



Wie sieht's in XML aus

```
<!--  
////////////////////////////////////  
//  
// Category (Subdivision Type) Individuals  
//  
////////////////////////////////////  
-->  
<owl:NamedIndividual rdf:about="&lcc-3166-2-bg;District">  
  <rdf:type rdf:resource="&lcc-cr;GeographicRegionKind"/>  
  <rdfs:label>district</rdfs:label>  
  <skos:definition>subdivision category district in the country of Bulgaria</skos:definition>  
  <rdfs:isDefinedBy rdf:resource="&lcc-3166-2-bg;"/>  
</owl:NamedIndividual>
```

Weltweit eindeutiger Name
URL-Form
<https://rdf.acolada.de/meineWelt>

Referenziert eine IRI
International Resource Identifier
Oben verwendete URL ist eine Form aber sehr gebräuchlich



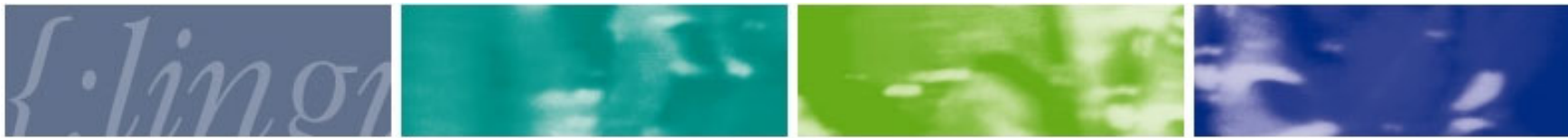
Wie sieht's in XML aus

```
<rdfs:Class rdf:about="http://iirds.tekom.de/iirds#Action">  
  <rdfs:label xml:lang="de">Handlung</rdfs:label>  
  <rdfs:label xml:lang="en">Action</rdfs:label>  
  <rdfs:subClassOf rdf:resource="http://iirds.tekom.de/iirds#FunctionalMetadata"/>  
  <iirds:description xml:lang="en">Use the relates-to-action property to create a reference from an information unit to an action.</iirds:description>  
  <rdfs:comment xml:lang="en">Class for atomic actions performed by users that are described in technical documentation.</rdfs:comment>  
</rdfs:Class>
```

Eigene Definition von
Eigenschaften/
Beziehungen

Mehrsprachige Definitionsmöglichkeit

IRI



Wo finde ich existierende RDF's

RDF VOCABULARIES

- ▼ RDF vocabularies for bibliographic and authority data
- ▼ RDF Element Sets
 - ▼ GND Ontology
 - ▼ DNB Metadata Terms
 - ▼ Agent Relationship Ontology
- ▼ RDF Value Vocabularies
 - ▼ GND Gender
 - ▼ GND Geographic Area Codes
 - ▼ GND Subject Categories
 - ▼ GND Type of Coordinates

Detailed Information

Classes

Administrative unit

URI	http://d-nb.info/standards/elementset/gnd#AdministrativeUnit
Label (German)	Verwaltungseinheit
Label (English)	Administrative unit
MARC 21 Equivalent	079 \$v=giv
Has superclass	Place or geographic name

Authority Resource

URI	http://d-nb.info/standards/elementset/gnd#AuthorityResource
Label (German)	Normdatenressource
Label (English)	Authority Resource
Has subclass	Conference or Event, Corporate Body, Family, Person, Place or geographic name, Subject heading, Work



Wo finde ich existierende RDF's

<https://id.loc.gov/> Library of Congress (USA)

<https://lod-cloud.net/> Linked open data cloud

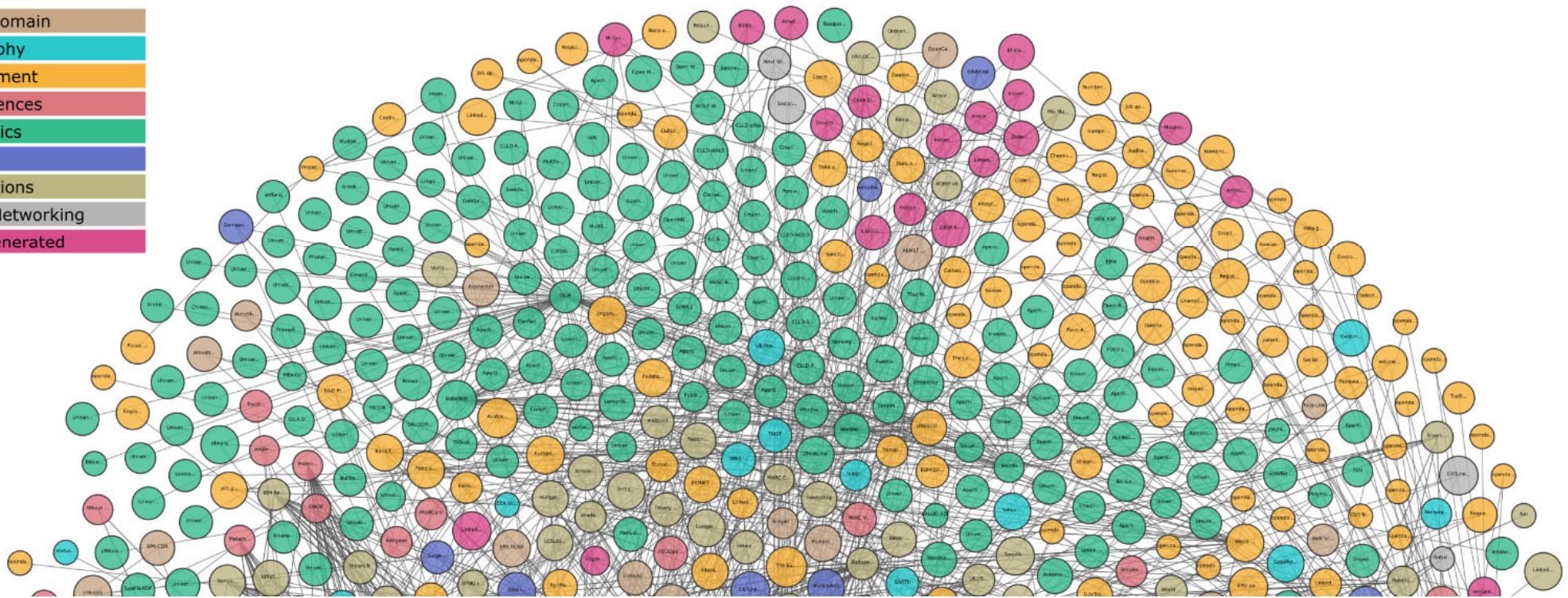
<https://www.dbpedia.org/> Dbpedia

....

The **Linked Open Data** Cloud

Legend

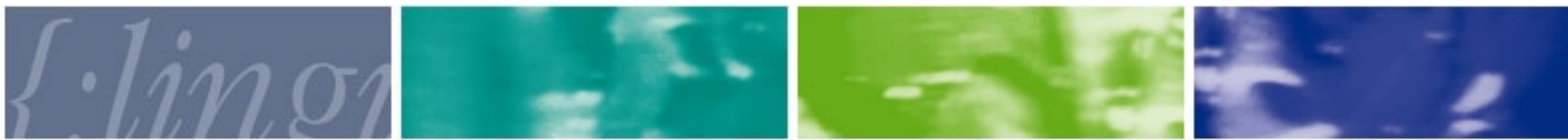
Cross Domain
Geography
Government
Life Sciences
Linguistics
Media
Publications
Social Networking
User Generated





Wofür?

- > Suche per Volltext wie in Google -> viele hundert Treffer
- > Zur Einengung -> mehr Suchbegriffe
- > Metadaten -> Facettensuche -> Bereichseingrenzung
 - > Oberbegriff/Unterbegriff -> Taxonomie -> Klassifiziere meine Dokumente
- > Semantic Web erlaubt Schlüsse zu ziehen = nicht manuell gesetzte Beziehungen abzuleiten
- > Viele Triple erstellen Graphen mit vielen Beziehungen
- > Suche kann sich diese Informationen zu nutze machen



Wofür?

- > Subjekt -> Prädikat -> Objekt
- > München -> ist eine -> Stadt
- > NamedIndividual -> isMemberOf -> Class
- > Analysiert man Texte dann kann man mit den gewonnenen Tripeln Schlüsse ziehen
 - > Welche Begriffe kommen vor
 - > Welche Beziehungen entstehen zu anderen Begriffen
 - > Viele solcher Beziehungen ermöglichen dann Abfragen, die sich nicht einfach aus dem Text ergeben
 - > Welche Städte in Bayern haben mehr als 1 Mio. Einwohner?
 - > 3 Triple (vereinfacht)
 - > München ist eine Stadt
 - > München liegt in Bayern
 - > München hat 1,5 Mio. Einwohner



Was bringt Linguistik?

- > Subjekt -> Prädikat -> Objekt
- > München -> ist eine -> Stadt
- > NamedIndividual -> isMemberOf -> Class
- > mit Linguistik kann man Texte analysieren und Triple erzeugen
- > man kann Begriffe finden -> Substantive
- > Man kann Komposita zerlegen -> Baumstamm nach Baum und Stamm
- > Man kann Adjektive/Verben erkennen -> hoher Baum -> Eigenschaft des Baums
- > Ziel: Verminderung des Aufwandes für Metadatengewinnung bzw. Triple-Erzeugung



Im Verlag?

- > Subjekt -> Prädikat -> Objekt
- > München -> hat -> ein Oberlandesgericht
- > OLG -> ist eine Abkürzung -> für Oberlandesgericht
- > München -> hat -> ein Landgericht
- > LG -> ist eine Abkürzung -> für Landgericht
- > Ein Landgericht -> ist eine -> Vorinstanz des Oberlandesgerichts
- > AZ -> ist eine Abkürzung -> für Aktenzeichen

- > Heute erging am LG München II das Urteil mit dem AZ 1234 ...
- > Heute erging am OLG München das Urteil mit AZ 1234 ...

- > Jetzt können wir daraus einiges Schlussfolgern bzw. auch dem Anwender etwas an zusätzlichen Informationen bieten ohne es je so definiert zu haben



Zusammenfassung

- > SemanticWeb ist nichts veraltetes
- > Linguistik und RDF können die Aufbereitung von Inhalten deutlich verbessern
- > Suche besser steuerbar aber auch related Content kann besser angebunden werden
- > Es ist nicht ohne Investition zu bekommen
- > eine Planung der Umsetzung ist wichtig, passende RDF's finden, eigene RDF's erstellen
- > was ist KI (künstlicher Intelligenz)
 - > sicher ein weiterer Baustein auf dem Weg die Inhalte besser zur erschließen
- > was bleibt ist ich muss mich mit den Fragen meiner Anwender beschäftigen, was wollen die wissen/finden