LEIBNIZ-INFORMATIONSZENTRUM
TECHNIK UND NATURWISSENSCHAFTEN
UNIVERSITÄTSBIBLIOTHEK

TIB

Linked Data & Ontologien – Ontologien

Dr. Anna Kasprzik
Hannover, 17. September 2018
3. VIVO-Workshop 2018

Leibniz Gemeinschaft

Vorbemerkung

Wissensorganisation wird schon
seit Jahrtausenden betrieben!

Und auch die Formale Semantik
ist mindestens 100 Jahre alt.

Das Semantic Web versucht lediglich, auf der Basis
von Linked Data diese Formalismen zu implementieren,
um semantische Information im Web
maschinenverarbeitbar zu machen.

Wir fassen zusammen:

Das Semantic Web besteht also

aus vernetzten Daten (Linked Data),

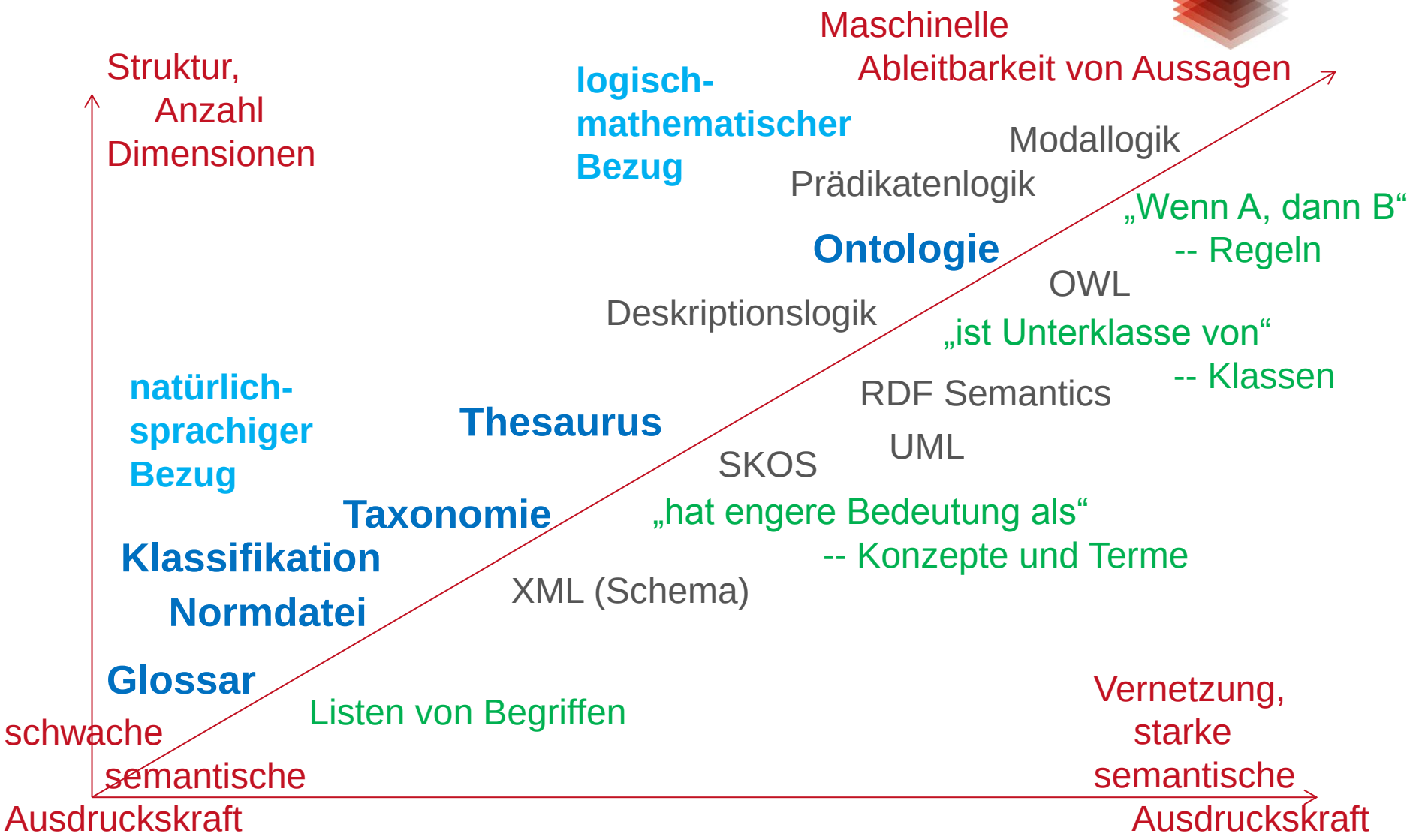
die mit semantischen Angaben annotiert sind.

Aus welchem Fundus kommen

diese semantischen Angaben?

Wissensorganisationssysteme: Das Spektrum

Wissensorganisationssysteme: Das Spektrum



Überblicksartikel: **Towards a Taxonomy of KOS : Dimensions for Classifying Knowledge Organization Systems.**
Souza, Renato R.; Tudhope, Douglas; Almeida, Mauricio B. In: Knowledge Organization, Vol. 39, No. 3, 2012, p. 179-192.

Glossar, Taxonomie, Thesaurus, Ontologie

- Glossar: Liste von Begriffen, gegebenfalls mit Definitionen und weiteren Informationen [linear]

- Taxonomie: Begriffe sind durch Ober-/Unterbegriffsrelationen hierarchisch geordnet [Baumstruktur]

- Thesaurus: (idealerweise) Trennung in Konzepte und Terme; weitere Relationen wie Synonymie und andere semantische Bezüge [Baumstruktur oder Netzstruktur, Polyhierarchie]

---------- natürlichsprachige Grenze ----------

- (Heavy-Weight-)Ontologie: Klassen, Relationen und Eigenschaften plus logische Schlussregeln [mengentheoretische Auffassung]

Wie annotiert man Ressourcen im Web

mit semantischen Angaben aus

Wissensorganisationssystemen?

Wie repräsentiert man

Wissensorganisationssysteme im Web?

Das Datenmodell RDF:
Resource Description Framework

Resource, Description, Framework

Resource

- alles, was durch einen URI
  eindeutig identifiziert und referenzierbar ist

Description

- über Relationen zu anderen Ressourcen und zu Datenwerten

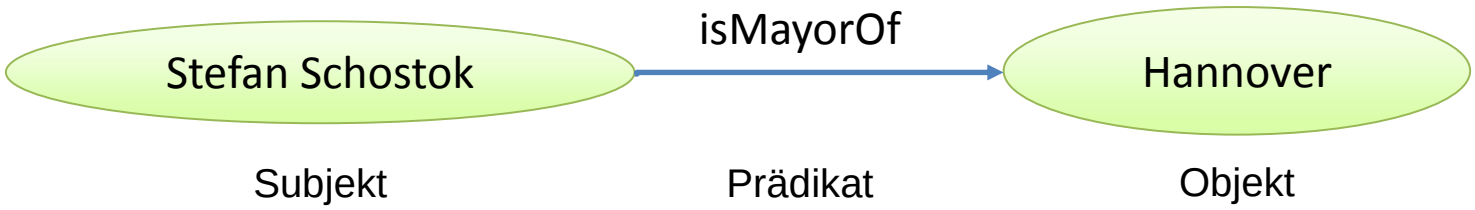
Framework

- basiert auf formalem Semantikmodell
- nutzt webbasierte Protokolle (URI, HTTP, XML, ...)

https://www.w3.org/TR/rdf11-primer/

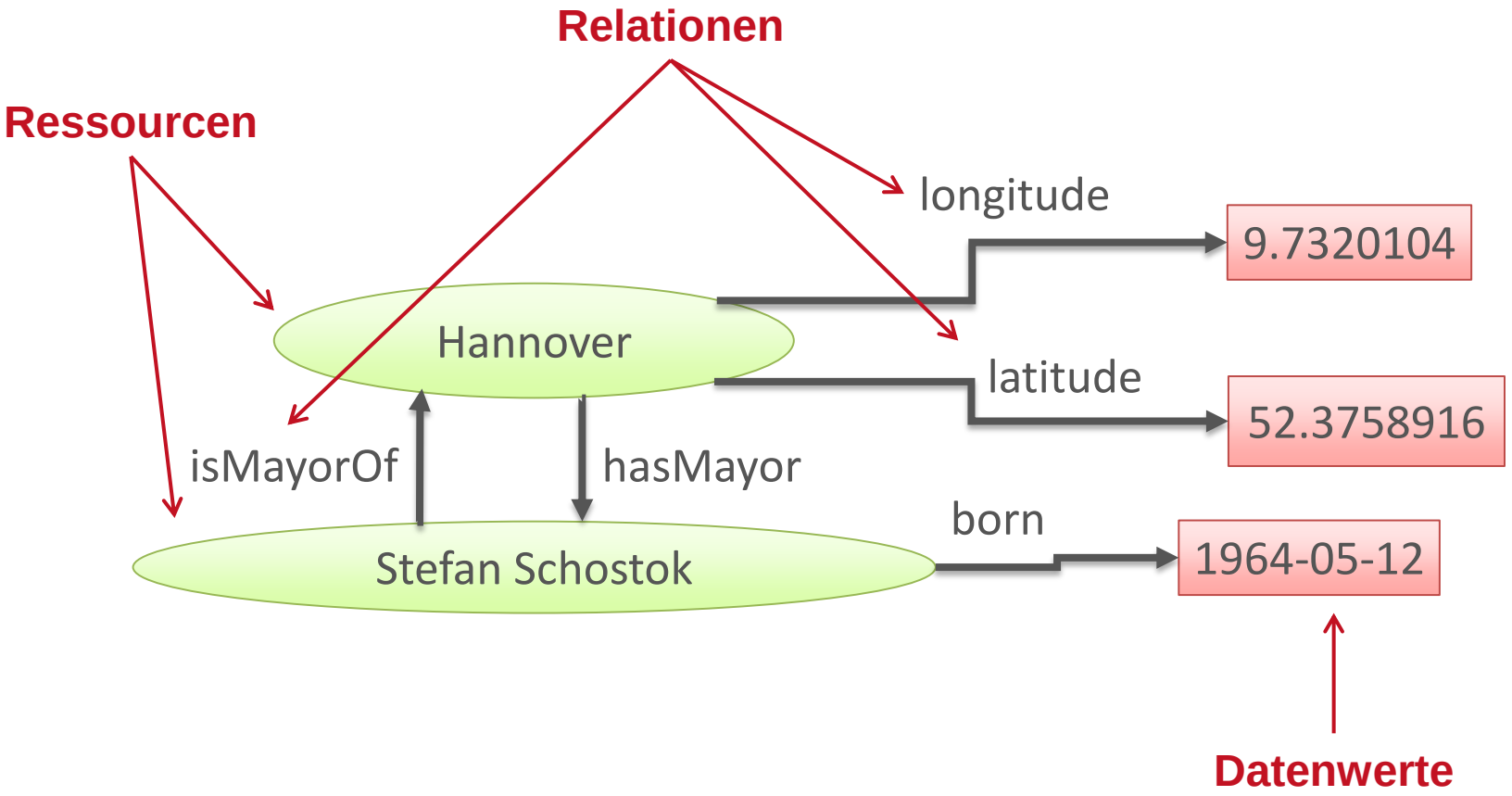
RDF-Tripel

Bestandteile eines RDF-Tripels:



- ... angelehnt an linguistische Kategorien (mehr oder weniger angemessen)
- Ressourcen – einschließlich Relationen! – werden durch URIs repräsentiert
- Weitere Bestandteile können sein:
  - Literale zur Repräsentation von Datenwerten als Objekte
  - „Blank Nodes“ als Platzhalter für Subjekte oder Objekte

Darstellung mehrerer RDF-Tripel als Graph



<http://dbpedia.org/resource/Hanover>
<http://dbpedia.org/resource/Stefan_Schostok>

RDF Syntax: Turtle

```
<http://dbpedia.org/resource/Hanover>
    <http://dbpedia.org/property/longitude>
        "9.7320104"^^<http://www.w3.org/2001/XMLSchema#float> .

<http://dbpedia.org/resource/Hanover>
    <http://dbpedia.org/property/latitude>
        "52.3758916"^^<http://www.w3.org/2001/XMLSchema#float> .

<http://dbpedia.org/resource/Hanover>
    <http://dbpedia.org/property/hasAreaCode>
        "0511" .

<http://dbpedia.org/resource/Hanover>
    <http://dbpedia.org/property/locatedIn>
        <http://dbpedia.org/resource/Lower_Saxony>.

<http://dbpedia.org/resource/Hanover>
    <http://dbpedia.org/property/hasMayor>
        <http://dbpedia.org/resource/Stefan_Schostok>.
```

RDF-Syntax: Turtle, mit Abkürzungen

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix ex: <http://example.org/property/> .
@prefix : <http://dbpedia.org/resource/> .

:Hanover
    ex:longitude   "9.7320104"^^xsd:float ;
    ex:latitude    "52.3758916"^^xsd:float ;
    ex:hasAreaCode "0511";
    ex:locatedIn   :Lower_Saxony, :Germany ;
    ex:hasMayor    :Stefan_Schostok .

:Lower_Saxony
    ex:locatedIn   :Germany .

:Stefan_Schostok
    ex:born        "1964-05-12"^^xsd:date ;
    ex:isMemberOf  :Social_Democratic_Party_of_Germany ;
    ex:isMayorOf   :Hanover .
```

RDF: Genügt das, um semantische Information zu beschreiben?

Angenommen, in einer RDF-Datei stehen die folgenden zwei Zeilen:

```
ex:Peter ex:istVaterVon ex:Paul .
ex:Paul ex:istVaterVon ex:Pip .
```

Gilt `ex:Peter ex:istVerwandterVon ex:Pip .` ?

Kann ich das einfach dazuschreiben?

Wie sieht das im folgenden Fall aus?

```
ex:Arx ex:sfgHdGh ex:Pof .
ex:Pof ex:sfgHdGh ex:Meh .
```

Gilt `ex:Arx ex:uHHfgsn ex:Meh .` ?

RDF Schema (RDFS)

- RDFS erlaubt die Spezifikation von Teilen der Semantik beliebiger RDF-Vokabulare und ist somit ein Metavokabular

- erlaubt Aussagen über generische Mengen von Individuen (Klassen), z.B. Verlage, Personen, Organisationen, etc.
- erlaubt Spezifikation logischer Zusammenhänge zwischen Individuen, Klassen und Relationen

  „Verlage sind Organisationen.“
  „Nur Personen schreiben Bücher.“

- Vorteil: Jede Software mit RDFS-Unterstützung interpretiert jedes mit RDFS definierte Vokabular korrekt

- Diese Funktionalität macht RDFS zu einer semantischen Beschreibungssprache (für sogenannte „*light-weight*“ Ontologien)

https://www.w3.org/TR/rdf-schema/

RDFS

Construct	Syntactic form	Description
Class (a class)	**C** `rdf:type rdfs:Class`	**C** (a resource) is an RDF class
Property (a class)	**P** `rdf:type rdf:Property`	**P** (a resource) is an RDF property
type (a property)	**I** `rdf:type` **C**	**I** (a resource) is an instance of **C** (a class)
subClassOf (a property)	**C1** `rdfs:subClassOf` **C2**	**C1** (a class) is a subclass of **C2** (a class)
subPropertyOf (a property)	**P1** `rdfs:subPropertyOf` **P2**	**P1** (a property) is a sub-property of **P2** (a p
domain (a property)	**P** `rdfs:domain` **C**	domain of **P** (a property) is **C** (a class)
range (a property)	**P** `rdfs:range` **C**	range of **P** (a property) is **C** (a class)

NOTE

The syntactic form (second column) is in a prefix notation which is discussed in more detail in Sec.
the constructs have two different prefixes (`rdf:` and `rdfs:`) is a somewhat annoying historical artefa
preserved for backward compatibility.

Jede Entität eines RDF-Modells ist eine Instanz der Klasse `rdfs:Resource`.

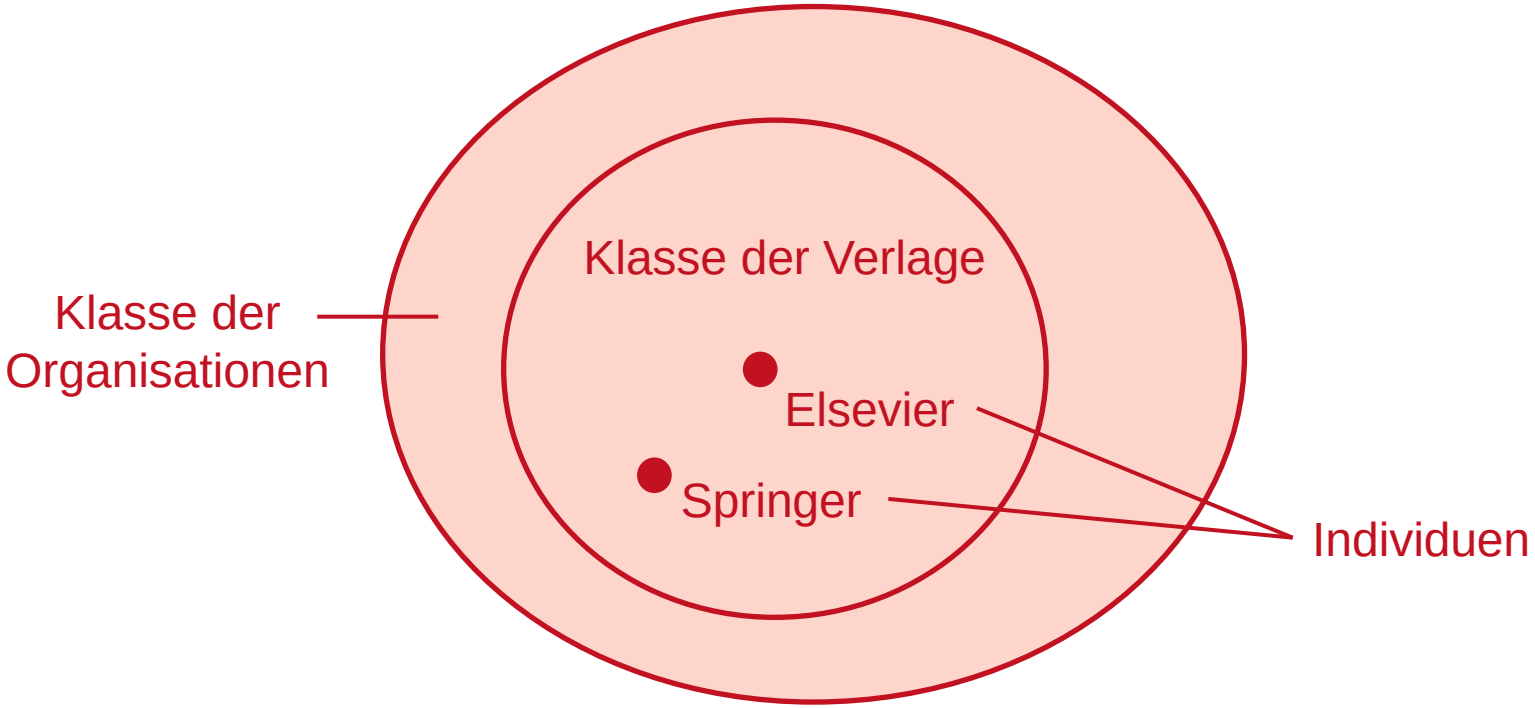
https://www.w3.org/TR/rdf11-primer/

Klassen und Individuen

```
ex:Organisation rdf:type rdfs:Class .
ex:Verlag rdf:type rdfs:Class .

ex:Verlag rdfs:subClassOf ex:Organisation .

ex:Elsevier     a     ex:Verlag .
ex:Springer     a     ex:Verlag .
```



Relationen in RDFS: Eigenschaften

```
ex:istVaterVon rdf:type rdf:Property .
ex:istVerwandterVon rdf:type rdf:Property .

ex:istVerwandterVon rdfs:domain ex:Person .
ex:istVerwandterVon rdfs:range ex:Person .

ex:istVaterVon rdfs:subPropertyOf ex:istVerwandterVon .
```

- Die Angabe einer Domain (Range) bedeutet, dass jede Entität, die als Subjekt (Objekt) der angebenen Relation auftaucht, automatisch (~ von einem Reasoner) dieser Klasse zugeordnet wird
- Alle Paare, die in einer Relation stehen, stehen automatisch auch in einer ihr mittels `rdfs:subPropertyOf` übergeordneten Relation

Literale, Datentypen, Strings mit Sprachtags

Die Klasse **`rdfs:Literal`** ist die Klasse von Literalen
wie z.B. Strings (Zeichenketten) und Integers (ganze Zahlen).

Die Klasse **`rdfs:Datatype`** ist die in RDF spezifizierte Klasse
von Datentypen. Sie ist mit XML Schema kompatibel.
Jede Instanz von `rdfs:Datatype` ist eine Unterklasse von `rdfs:Literal`.

```
ex:Peter ex:hatAnzahlKinder "5"^^xsd:integer .
```

Die Klasse **`rdf:langString`** ist die Klasse von Zeichenketten mit
Sprachtags. `rdf:langString` ist eine Instanz von `rdfs:Datatype`.

```
ex:DeAardappeleters
   ex:hatTitel
     "Die Kartoffelesser"@de .
ex:DeAardappeleters
   ex:hatTitel
     "The Potato Eaters"@en .
```

Das Format für Thesauri im Semantic Web: SKOS

Thesauri im Semantic Web: SKOS

- SKOS: Simple Knowledge Organisation System
- Zur Repräsentation von kontrollierten Vokabularen;
  geeignet, um weniger formale, natürlichsprachige Beziehungen auszudrücken, z.B. zur Begriffsklärung und Disambiguierung
- Basis: RDF/RDFS (Tripel)
- Einfach, flexibel, erweiterbar, maschinenlesbar
  → Veröffentlichung im World Wide Web
  → ermöglicht Austausch zwischen Softwareanwendungen
- Pragmatischer Kompromiss zur stark logisch ausgerichteten Web Ontology Language OWL

http://www.w3.org/TR/skos-primer

Semantische Beziehungen in SKOS

- skos:broader

ex:Säugetiere rdf:type skos:Concept;
skos:prefLabel „Säugetiere“@dt;
skos:**broader** ex:Tiere.

- skos:narrower

ex:Tiere rdf:type skos:Concept;
skos:prefLabel „Tiere“@dt;
skos:**narrower** ex:Säugetiere.

- skos:related

ex:Vögel rdf:type skos:Concept;
skos:prefLabel „Vögel“@dt
skos:**related** ex:Ornithologie.

SKOS: Label

- skos:prefLabel

ex:Tiere rdf:type skos:Concept;
skos:**prefLabel** „Tiere“**@dt**;
skos:**prefLabel** „animals“**@en**.

- skos:altLabel

ex:Tiere rdf:type skos:Concept;
skos:prefLabel „animals“@en;
skos:**altLabel** „creatures“@en.

- skos:hiddenLabel

ex:Tiere rdf:type skos:Concept;
skos:prefLabel „Tiere“@dt;
skos:**hiddenLabel** „Tire“@dt.

SKOS: Zwei standardisierte Erweiterungen

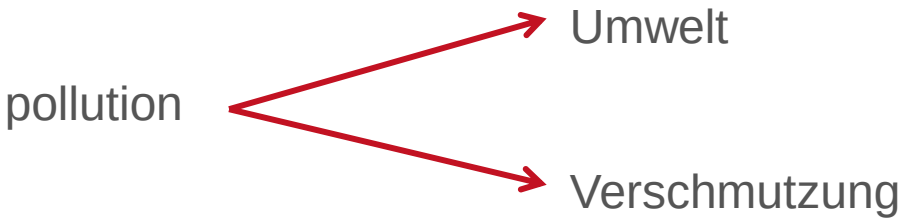
- SKOS-XL:
  Label als
  eigene Klasse!
  Label können so
  mit Metadaten
  versehen werden.

```
ex:C_Tiere rdf:type skos:Concept;
      skosxl:prefLabel ex:Tiere;
      skosxl:prefLabel ex:Animals.

ex:Tiere rdf:type skosxl:Label;
      skosxl:literalForm „Tiere“@dt.

ex:Animals rdf:type skosxl:Label;
      skosxl:literalForm „animals“@en.
```

- iso-thes: Aufbauend auf dem neuen ISO-Standard 25964-2
  zur Interoperabilität von Thesauri;
  erlaubt es u.A., komplexere Relationen auszudrücken, z.B.:



„zusammengesetzte Äquivalenz“

SKOS: Fazit

“The aim of SKOS is not to replace original conceptual vocabularies in their initial context of use, but to allow them to be ported to a shared space, based on a simplified model, enabling wider re-use and better interoperability.“

SKOS Simple Knowledge Organization System Primer
W3C Working Draft 29 August 2008 (http://www.w3.org/TR/skos-primer)

Die Web Ontology Language: OWL

Bestandteile/Grundbegriffe

- **Axiome:**
  die grundlegenden Aussagen, die eine OWL-Ontologie trifft

  “HUND ist eine Klasse in dieser Ontologie.”

  “Fifi gehört zur Klasse der Hunde in dieser Ontologie.”

- **Entitäten:** Elemente, die auf Objekte in der echten Welt hinweisen

  `ex:Fifi`

- **Ausdrücke:** Kombinationen von Entitäten, die aus grundlegenden Aussagen komplexe Aussagen formen

  “Die Großelternrelation lässt sich durch die zweifache Verkettung der Elternrelation beschreiben.”

https://www.w3.org/TR/owl2-primer/

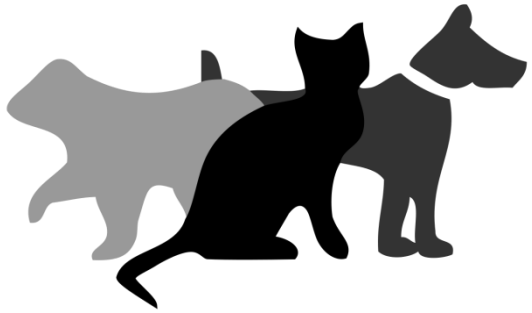
Object Properties und Data Properties

- Object Properties spezifizieren Relationen zwischen Entitäten

- Data Properties spezifizieren Relationen zwischen Entitäten und Datenwerten

Komplexere Axiome

- z.B. Disjunktheit:
  „Keine Entität ist zugleich ein Hund und eine Katze.“



- oder Zahlenbeschränkungen:
  „John hat höchstens vier Kinder.“

```
:John rdf:type [ rdf:type owl:Restriction ;
      owl:maxQualifiedCardinality
            "4"^^xsd:nonNegativeInteger ;
      owl:onProperty  :hasChild ;
      owl:onClass  :Parent ] .
```

Praxisübung

Spielregeln

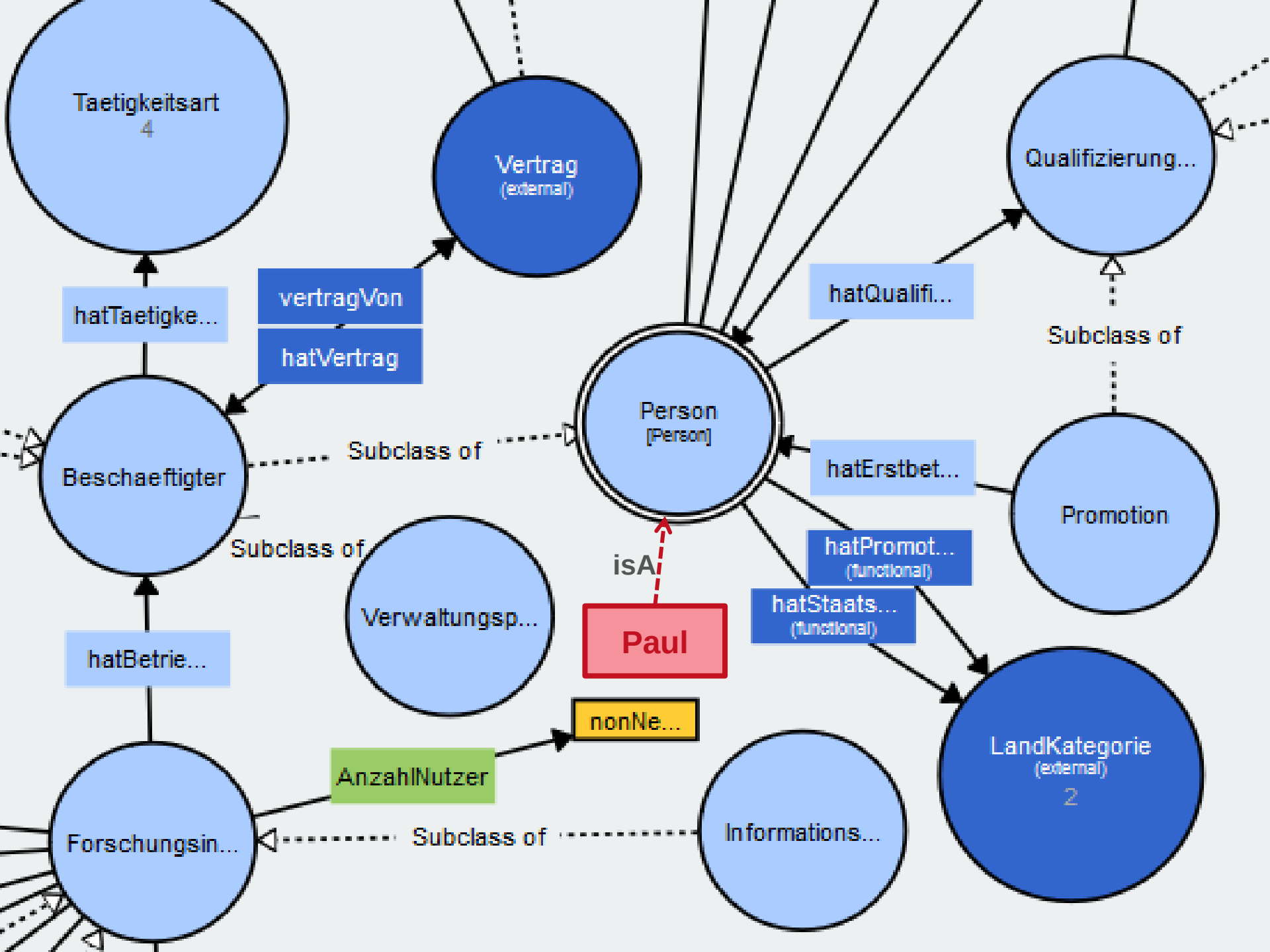
Bilden Sie vier Gruppen.

Bearbeiten Sie den vorgeschlagenen Themenbereich (Mikrowelle) oder suchen Sie sich einen eigenen. Überlegen Sie:

- Welche Klassen modellieren den Zielbereich am besten?
- Welche Relationen modellieren den Zielbereich am besten? Zwischen Klassen? Zu Datenwerten?
- Welche zusätzlichen Sachverhalte möchten Sie abbilden, funktioniert das mit den von Ihnen definierten Klassen und Relationen?

Zeichnen Sie Ihre Ontologie als Graphen auf ein DIN-A4-Blatt, übertragen Sie es dann nach Fertigstellung auf Ihre Stellwand

Dokumentieren Sie, auf welche Probleme und Fragen Sie beim Modellieren gestoßen sind.


Taetigkeitsart
4
Vertrag
(external)
Qualifizierung...
hatTaetigke...
vertragVon
hatVertrag
hatQualifi...
Subclass of
Person
[Person]
Beschaeftigter
Subclass of
hatErstbet...
Promotion
Subclass of
isA
hatPromot...
(functional)
Verwaltungsp...
Paul
hatStaats...
(functional)
hatBetrie...
nonNe...
LandKategorie
(external)
2
AnzahlNutzer
Forschungsin...
Subclass of
Informations...

LEIBNIZ-INFORMATIONSZENTRUM
TECHNIK UND NATURWISSENSCHAFTEN
UNIVERSITÄTSBIBLIOTHEK

TIB

Danke.

Kontaktdaten
Dr. Anna Kasprzik
T 0511 762-14219, anna.kasprzik@gmx.de

Leibniz Gemeinschaft