

# **Developing an Ontology for Academic Disciplines**

---

Martin Stricker

Hermann von Helmholtz-Zentrum für Kulturtechnik

Humboldt-Universität zu Berlin

15 October 2008

# Hermann von Helmholtz-Zentrum für Kulturtechnik

---



- An interdisciplinary research center at Humboldt University, Berlin
- „Kulturtechnik“ – the study of scientific and cultural shifts in the context of technical innovations
- Focal point (among other things) is communication in and of science and especially the role of „objects“ and scientific collections therein



HERMANN VON HELMHOLTZ-ZENTRUM FÜR KULTURTECHNIK  
Katalog der wissenschaftlichen Sammlungen der Humboldt-Universität zu Berlin (Pilotprojekt)



Wissenschaftliche Sammlungen > Suche

## Suche nach Sammlungsobjekten

Auf dieser Seite können Sie komplexe Suchabfragen zusammenfügen: Im ersten Feld grenzen Sie per Dropdownmenü den Suchbereich ein. "Suchfeld hinzufügen" erstellt einen Eingabebereich für den Suchbegriff. Auf diese Art können beliebig viele Zeilen hinzugefügt werden. Ist Ihre Abfrage fertig zusammengesetzt, wird sie mit "Suche" gestartet und Sie erhalten eine Liste der Ergebnisse.

Combine with

[Previous Page](#) 4 / 4



Exlibris, Rudolf Virchow



Fotografie, Der Fußkünstler Unthan mit Geige



Fotografie, Präparierübungen



Foto, Hans Virchow beim Präparieren



## Navigation

[Sammlungsobjekte](#)

[Thesaurus](#)

[Sammlungen](#)

[Erweiterte Suche](#)

[Die Geschichte der Sammlungen](#)

[Projektdefinition](#)

[Links und Tipps](#)

[Objekte des Monats](#)

[Impressum & Kontakt](#)

## Schnellsuche



## Objekt des Monats 3/2008



Wenkersätze in einer hessischen Mundart

das unermüdliche und  
ws (1821-1902) von  
er zweiten Hälfte des 19.  
) bestehende Berliner  
Urgeschichte besaß  
und von Stiftungsmitteln als  
ber persönliche  
Jahre nach Virchows Tod  
thropologischen  
1868-1931), späterer  
ernahm den Auftrag zu ihrer  
s (1854-1924)  
ernach ihre definitive,  
Nummer. Der gleichzeitig in  
er Wirren des Zweiten  
nlungszugänge gab es  
e Zeit zwischen 1930 und  
dig im Dunkeln. Ab 1964  
tizierungsarbeiten. Die  
de des Jahres 2004 durch  
versität zu Berlin  
re Übernahme durch das  
lung verbleibt jedoch  
Anthropologischen Instituts



## Objektgruppen Auswahl

Im Index „Objektgruppen“ ermöglichen innerhalb der Domänen *res naturalia* und *res artificialia* Beschreibungsaspekte den Zugang zu relevanten Gruppen von Objekten aus unterschiedlichen Perspektiven · Dokumentation (PDF)

Domänen und Beschreibungsaspekte	Auswählen	Auswahl
<ul style="list-style-type: none"> <li>⊕ <b>res naturalia</b></li> <li>⊖ <b>res artificialia</b> <ul style="list-style-type: none"> <li>Kunstgattung</li> <li>Funktion/Form</li> <li>Verwendungszweck</li> <li>Lehrmaterial/Demonstrationsmaterial</li> <li>Medium</li> <li>Forschungsbereich</li> <li>Material</li> <li>Technik/Herstellungsart</li> <li>Darstellungsthema</li> <li>Entstehungszeit</li> <li>Herkunftsregion</li> <li>Zusatzinformation</li> </ul> </li> </ul>	<b>res artificialia</b> <b>Kunstgattung</b> <ul style="list-style-type: none"> <li>Angewandte Kunst</li> <li>⊖ Architektur           <ul style="list-style-type: none"> <li>Architekturelemente</li> <li>Architekturzeichnungen/Architekturpläne</li> </ul> </li> <li>⊖ Bildende Kunst           <ul style="list-style-type: none"> <li>Buchkunst</li> <li>Fotografie</li> <li>⊖ Grafik               <ul style="list-style-type: none"> <li>Druckgrafik</li> <li>Zeichnung</li> <li>Zeichnungen</li> </ul> </li> <li>Installation</li> <li>⊖ Malerei               <ul style="list-style-type: none"> <li>Wandmalerei</li> </ul> </li> <li>⊖ Plastik/Skulptur               <ul style="list-style-type: none"> <li>Kleinplastik</li> <li>Reliefs</li> </ul> </li> <li>Textilkunst</li> </ul> </li> <li>⊖ Darstellende Kunst           <ul style="list-style-type: none"> <li>Bühnenbilder</li> <li>Requisiten</li> </ul> </li> <li>Design</li> <li>Musik</li> </ul>	<b>res artificialia</b> Angewandte Kunst Kunstgattung <span>✕</span>  <b>3 Sammlungen</b> · Aufrufen

## Objektgruppen Auswahl

Im Index „Objektgruppen“ ermöglichen innerhalb der Domänen *res naturalia* und *res artificialia* Beschreibungsaspekte den Zugang zu relevanten Gruppen von Objekten aus unterschiedlichen Perspektiven · Dokumentation (PDF)

Domänen und Beschreibungsaspekte	Auswählen	Auswahl
<input checked="" type="checkbox"/> <i>res naturalia</i>	<i>res naturalia</i>	<i>res naturalia</i>
<input checked="" type="checkbox"/> <i>res artificialia</i>	<i>res artificialia</i>	<i>res artificialia</i>
<input type="checkbox"/> Kunstgattung	Kunstgattung	Kunstgattung
<input type="checkbox"/> Funktion/Fo		
<input type="checkbox"/> Verwendung		
<input type="checkbox"/> Lehrmaterial		
<input type="checkbox"/> Medium		
<input type="checkbox"/> Forschungs		
<input type="checkbox"/> Material		
<input type="checkbox"/> Technik/Her		
<input type="checkbox"/> Darstellungs		
<input type="checkbox"/> Entstehung		
<input type="checkbox"/> Herkunftsre		
<input type="checkbox"/> Zusatzinform		

## Thesaurus

Der hier verwendete Thesaurus operiert transdisziplinär und polyhierarchisch. Sie können direkt im Suchfeld

[Sammlungen](#) · [Suche](#) · [Personen](#) · [Objekte](#) · [Bibliographie](#) · [Dokumentation](#) · [Weblog](#)

Alle Sammlungsschwerpunkte (Polyhierarchischer Thesaurus)

\*Aufgelöste, ausgelagerte, unbekannt verbliebene oder verlorene Sammlungen

Hierarchisch · Alphabetisch: [A](#) · [B](#) · [C](#) · [D](#) · [E](#) · [F](#) · [G](#) · [H](#) · [I](#) · [J](#) · [K](#) · [L](#) · [M](#) · [N](#) · [O](#) · [P](#) · [R](#) · [S](#) · [T](#) · [U](#) · [V](#) · [W](#) · [Z](#)

### M

Markscheidewesen  
Maschinenwesen  
Materialkunde  
Mathematik  
Mediävistik  
Medizin  
Medizingeschichte  
Medizinische Anthropologie  
Medizintechnik  
Metallurgie  
Mikrobiologie  
Mineralogie  
Musikethnologie  
Musikwissenschaft

### Mikrobiologie (2)

Oberbegriff  
BT Biologie

Unterbegriff  
NT Phylogenie

Benutzt für  
UF Microbiology

<a href="#">&gt; Anthropologie 12</a>	<a href="#">&gt; Informatik</a>	<a href="#">&gt; Physiologie 208</a>
<a href="#">&gt; Archäologie 179</a>	<a href="#">K &gt; Kartographie 8</a>	<a href="#">&gt; Politologie 14</a>
<a href="#">&gt; Astronomie 12</a>	<a href="#">&gt; Kulturwissenschaft 1</a>	<a href="#">&gt; Psychologie 4</a>
<a href="#">B &gt; Bakteriologie 11</a>	<a href="#">&gt; Kunstgeschichte 664</a>	<a href="#">R &gt; Radiologie 1</a>
<a href="#">&gt; Bibliothekswissenschaft 2</a>	<a href="#">L &gt; Linguistik 57</a>	

# Limitations of common Knowledge Organisation Systems and implementations

---

- Knowledge engineering level: Limited expressivity
- Implementation level: Each vocabulary has its own structure on the application level (application code, encoding or database schema)
- Interoperability level: Sharing or synchronizing information or data requires for each instance dedicated application code, and very often dedicated user interfaces for manual mapping

# Solution for the domain of Science and Humanities

---

- An "ontology" which respects the qualities of traditional knowledge organisation systems (poly-/hierarchies, synonyms), enhanced by additional (syntagmatic) relation types and logical inferencing potentials
- The model enables the description of disciplinary characteristics and interdisciplinary connections via properties/relations of typed significant constituents
- Ontology encoded in RDF-based ontology language OWL (interoperability for information exchange, chance of reusable application code)

# Design characteristics

---

- Terminological (linguistic) approach (currently rather low degree of formalization, lightweight inferencing)
- Basic usage shall support intuitive browsing and navigation, serendipitous discovery, assisted search (including entry points for more complex rule-based operations and representations)
- Main principle: As simple and intuitive as possible, as complex as necessary



# Usage scenarios

---

- Indexing, browsing, (assisted or faceted) searching of collections and databases which contain data on a supra- or multidisciplinary level
  - Bibliographic information (library catalogs), research databases, university information systems, outside representation ...
- Integration of legacy vocabularies possibly from heterogenous sources
- Wider integration within the Semantic Web / Linked Data structures

# Project team

---

## **Lead**

Dr. Cornelia Weber, General Manager of Helmholtz-Zentrum

## **Principal developers**

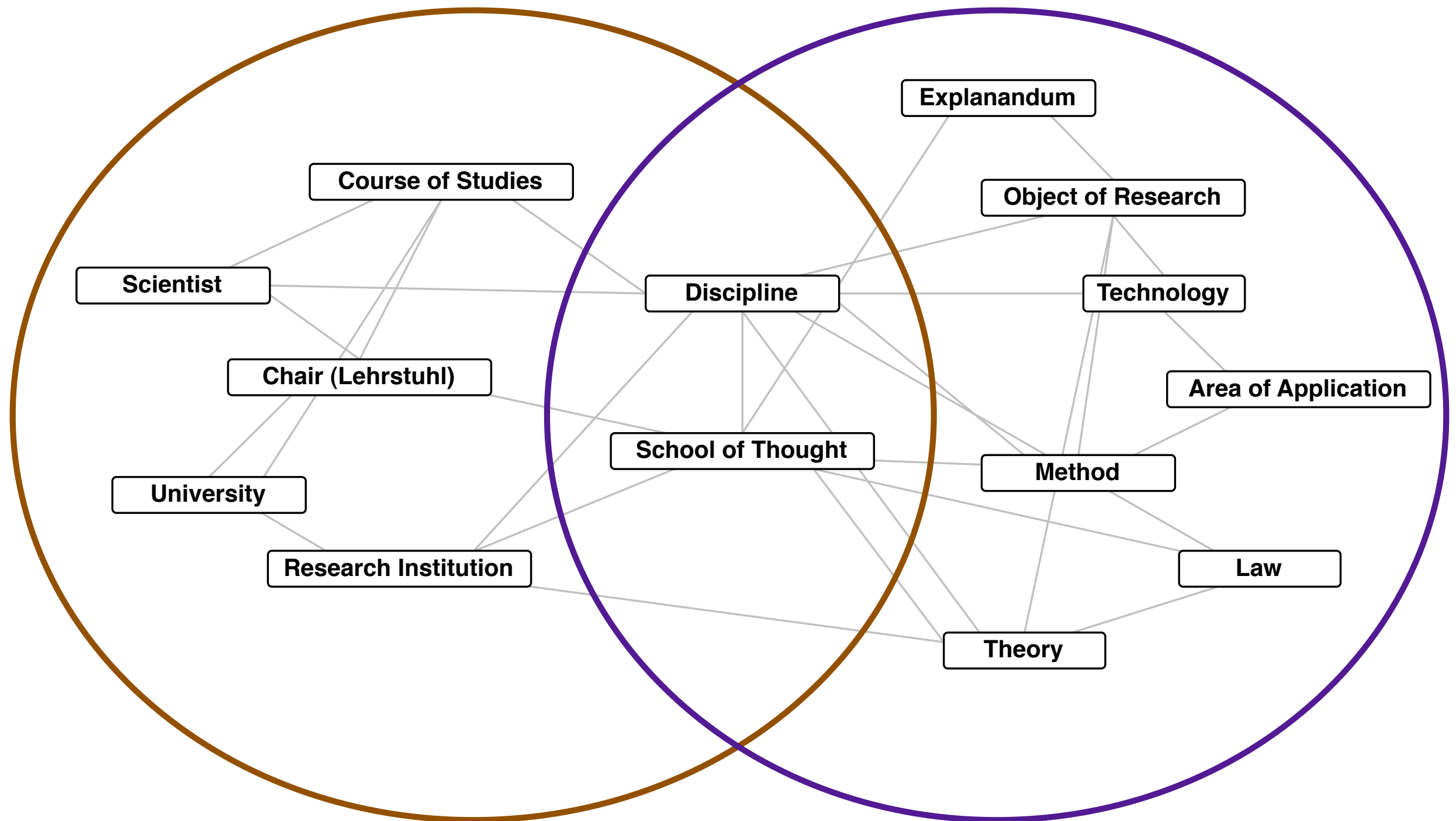
Dipl.-Phil. Information Specialist Elena Semenova  
*Concept, modelling*

Martin Stricker

*Semantic technologies & standards, software development*

# Top-Level Classes

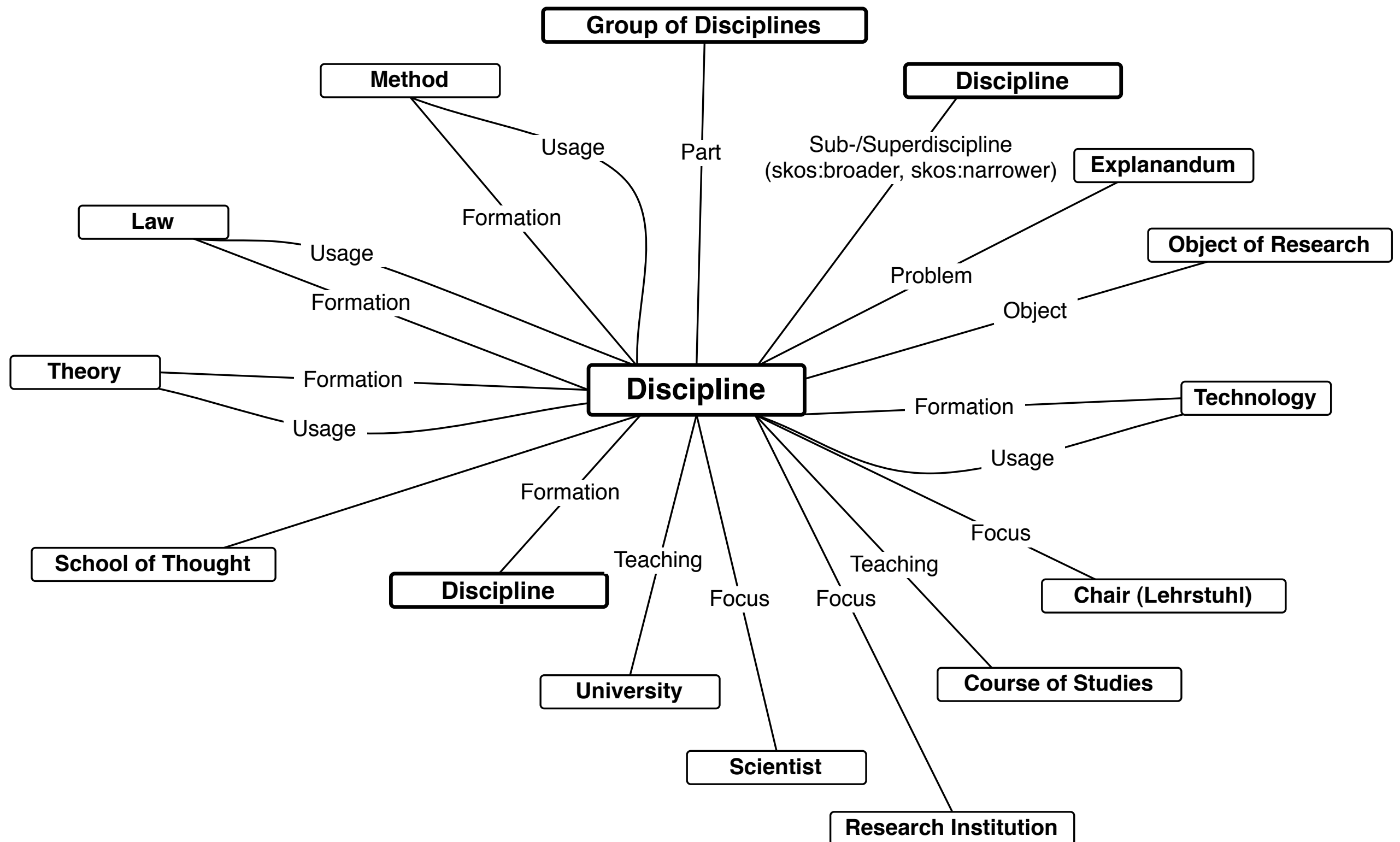
---



Institutional (social) subdomain

Epistemic subdomain

# Class properties: Discipline





# Specifications

---

- OWL 1 (recommendation), SKOS (latest working draft)
- Usage of class descriptions (local property restrictions) / property descriptions for consistency enforcing & checking
- Usage of SKOS:
  - Hierarchical relations
  - (Natural language) names or labels (`skos:prefLabel`, `skos:altLabel`, `skos:hiddenLabel`)

# Challenge: Communication

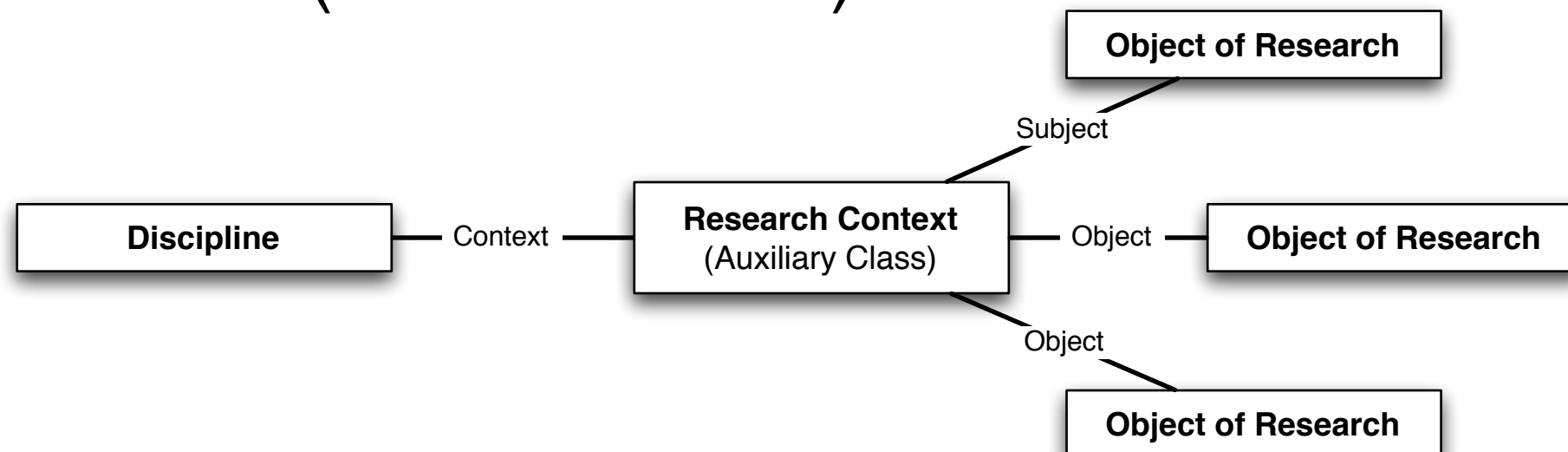
---

- Interdisciplinary communication is hard (especially in a project about interdisciplinary communication)
- Terms and meaning:
  - *Class, Instance, Individual*
  - *Concept, Resource, (Topic, Association)*
  - *Name, URI, IRI, Label*
  - plus the finer points of <xml>, Namespaces, First Order Logic, UML graphs, Open-World Assumption, Unique-Name Assumption

# Challenge: Contextual statements

---

- RDF only allows binary relations or properties (relations between two resources/individuals)
- How do you model this: *"In narratology, temporal and spatial qualities are aspects in the study of narratives"*
- We had to „reify“ relations/properties as auxiliary resources (class/individual)



# Challenge: Vocabulary Mapping

---

- How to integrate other (legacy) vocabularies and map their terms to concepts of the ontology?
  - *Schlagwortnormdatei* (SWD), *DDC Deutsch* (via CrissCross SWD-DDC Mapping), *Regensburger Verbundklassifikation* (RVK), *Personennamendatei* (PND), *Gemeinsame Körperschaftsdatei* (GKD)
- Missing namespaces and RDF representation
- Legal problems (copyright)
- For the actual mapping, we created auxiliary mapping classes in order to be able to annotate mapping relations (SKOS Mapping Properties can be auto-generated via rules/reasoner)



# Other challenges

---

- Tools and software for ontology editing: Reliability, performance, permanent „beta“, ...
- Semantic (re-)factoring/ further formalization, inferencing
- Internationalization
- Modelling and describing timelines and historical processes
- Implementation strategies for modular remote access: Client/server-architecture, interfaces, API, protocols, ...

# Further development: Core Ontology

---

- Refine and refactor model
  - Explore formalization options, logical rule-based reasoning
- Build a core vocabulary for a description of the German Science and Humanities disciplines (stable/long-term)
  - This in itself could trigger an application, a kind of modern interactive lexicon, tapping into other resources (Linked Data, Wikipedia/dbpedia)
  - The core vocabulary can be used as reference/background for applications which index specialised data or metadata for various purposes

# Further development: Projects and implementations

---

- Cooperate with partners who have a need or see an option to develop, use or integrate our model & vocabulary
- Consulting or development:
  - Model extension
  - Vocabulary development
  - Integration of legacy data
  - Software development (especially modular or service-based approaches)