

Wissensrepräsentation

Vorlesung

Sommersemester 2008

10. Sitzung

Dozent

Nino Simunic M.A.

Computerlinguistik, Campus DU

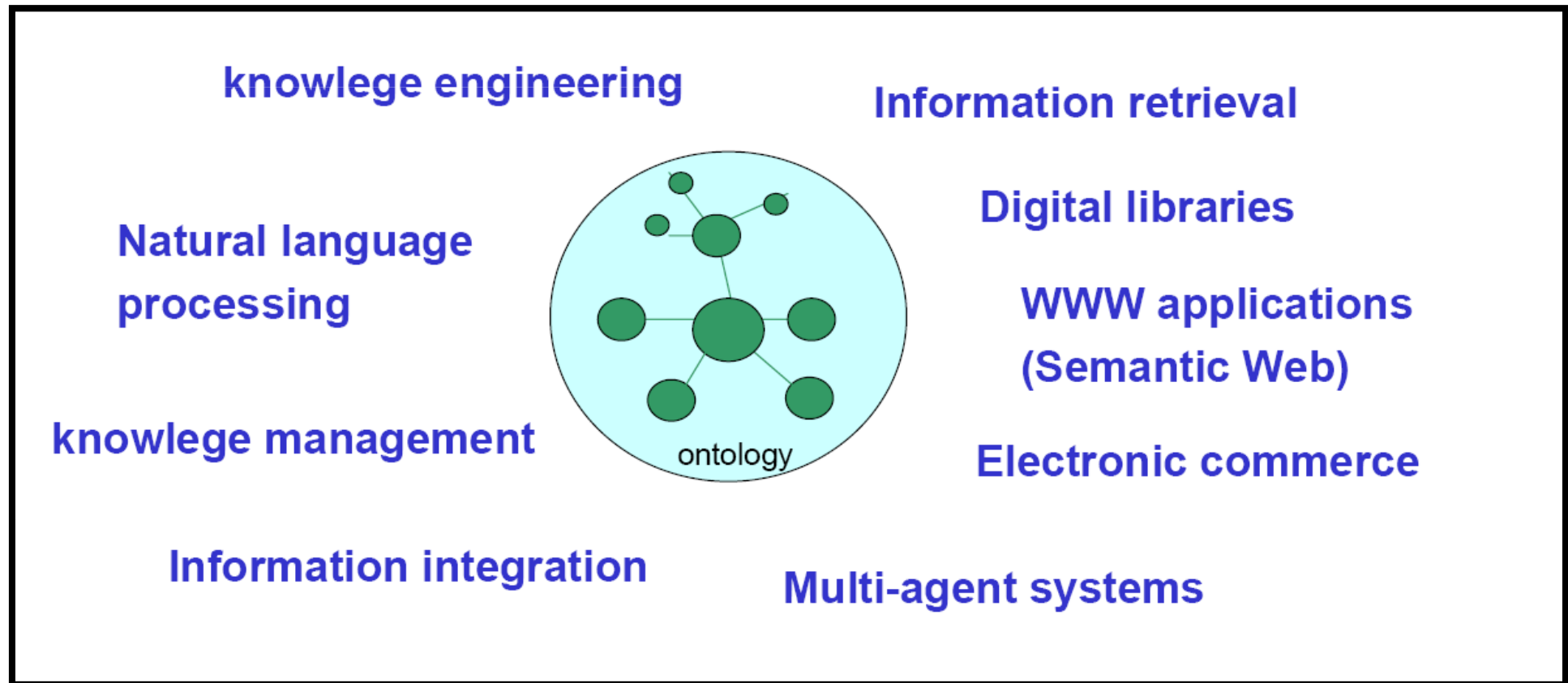
- **Ontologien: Einführung, Überblick**

Überblick: Ontologien in der Informatik

- Deklarative Repräsentation von Wissen in einer Interessengemeinschaft
- Wichtige Bestandteile des Formalismus
 - Konzepte, Relationen
 - Axiome, Instanzen
 - Basis: Logik
- Repräsentationssprachen zur Beschreibung
 - Semantic Web Sprachen (RDF, OWL, DAML, ...)
- Granularität, Aussagekraft, Ausdrucksstärke variabel je nach Typ
- Trennung zwischen Verarbeitung (Inferenzmaschine) und Beschreibung (Ontologie)
- Definition (kurz aber präzise):
 - »[S]hared and common understanding of a domain that can be communicated between people and application systems.«

[DAVIS:2003]

Ontologien: Einsatzgebiete



[MAEDCHE:2000]

Industrie, Wirtschaft, Unternehmen: Wissensmanagement

Lange Tradition:

- Materieller Besitz: Geld, Personen, Ressourcen, Höhe der eigenen Wolkenkratzer
- Kriterium für Unternehmenserfolg und Wettbewerb

»[I]t has been increasingly recognized **that in the post-industrial era, an organization's success is more dependent on its intellectual assets** than on the value of its physical resources.«

»The **competitiveness** of many companies depends heavily **on how they exploit their corporate knowledge and memory. [...]**«

[DAVIS:2003]

→ Wissensmanagement-Boom

Corporate Memory Management

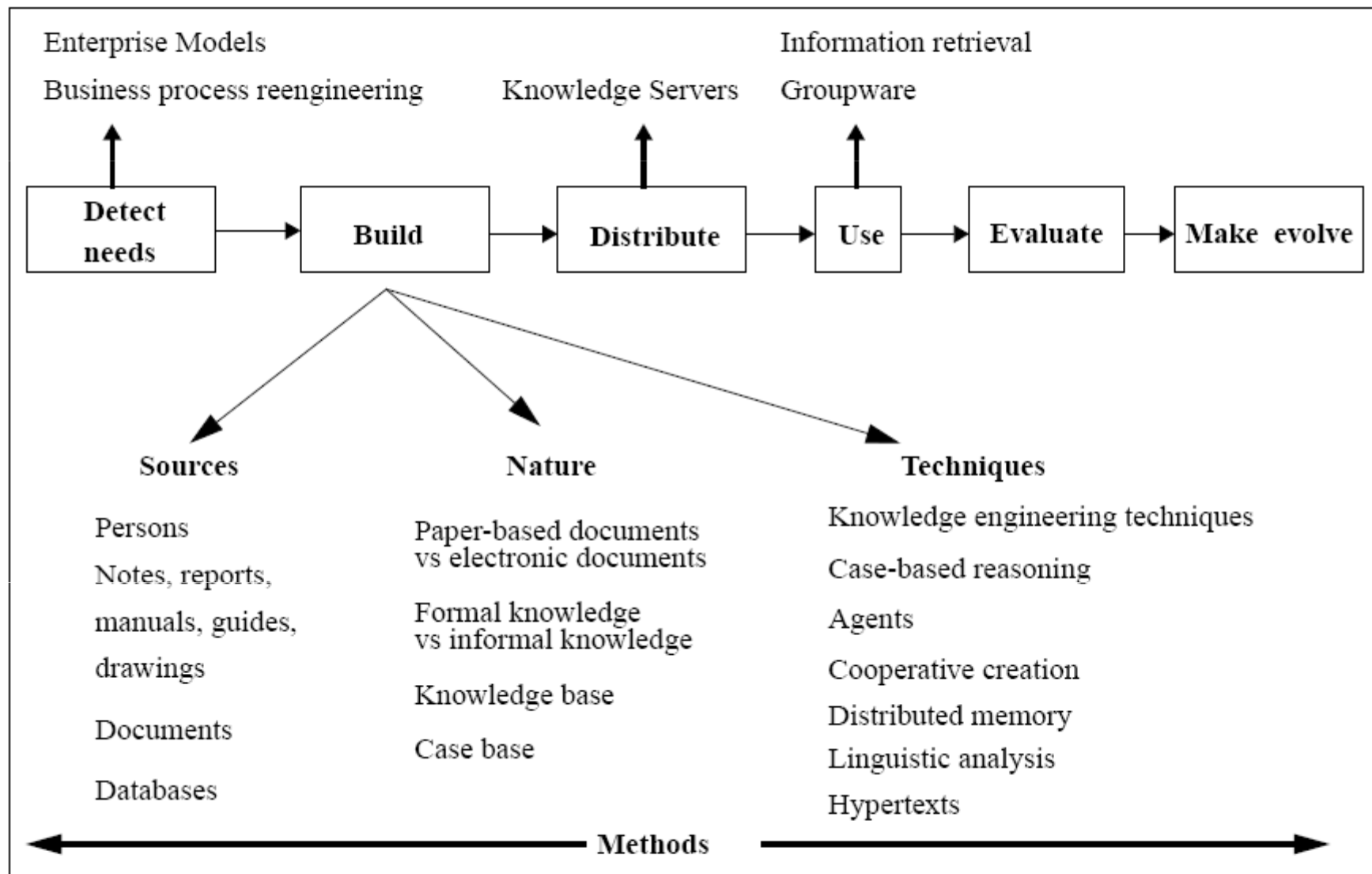


Fig.1: Corporate Memory Management

Ontologien fürs WWW

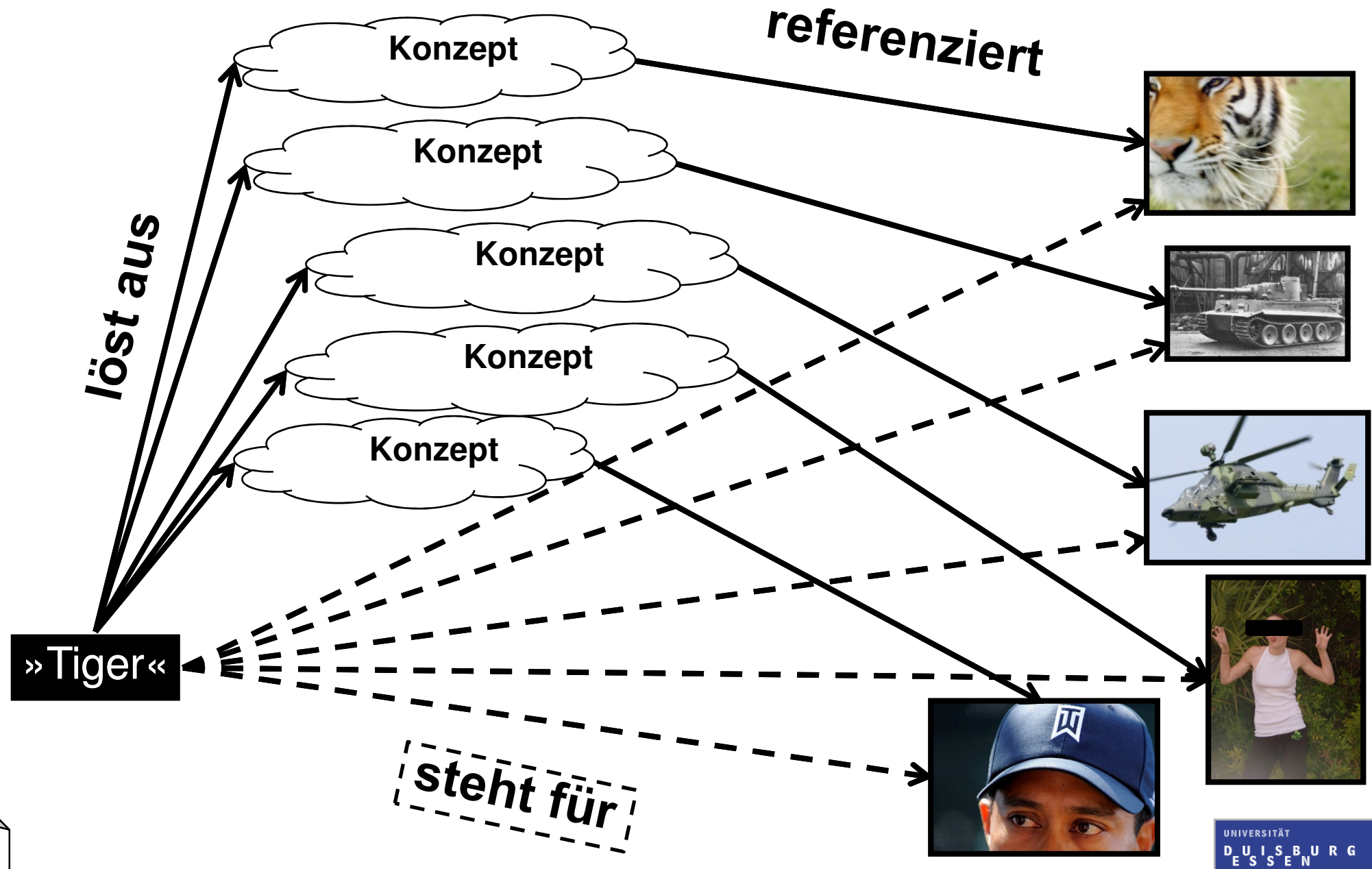
- »[...**Several**] **billion documents on the** [...]**WWW**, [...] used by more than **300 million users globally**, and **millions more** pages on **corporate intranets** [...make it] increasingly **difficult to find, organize, access** and maintain the information required by users.
[...] **Semantic Web** [...] provides **enhanced information access** based on the exploitation of **machine-processable meta-data** [...]«

[DAVIS:2003]

Einsatz von Ontologien

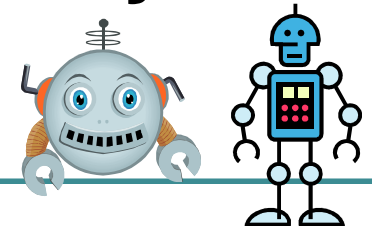
- Unterstützung der Kommunikation zwischen Personen, Organisationen, Prozessen, Systemen, Agenten
 - Durch einheitliche Konzeptionalisierung und Sprache
- Mediator-Funktion bei heterogenen Systemen / Schemata
 - Interoperabilität
 - Wissensaustausch & Wiederverwendung

- Symbol → mehrere Bedeutungen
- Indirekte Beziehung zwischen Symbol und Realität



Basis für erfolgreiche Kommunikation

»People can't share knowledge if they don't speak a common language«
[Thomas Davenport (1993)]



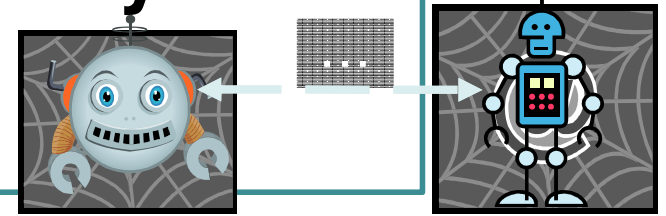
Impliziert hier:

Schnittmenge zwischen Kommunikationspartnern in

- Syntax,
- Semantik,
- Pragmatik/Kontext, (...)

Sonst: Erfolgreiche Kommunikation schlecht bis unmöglich.

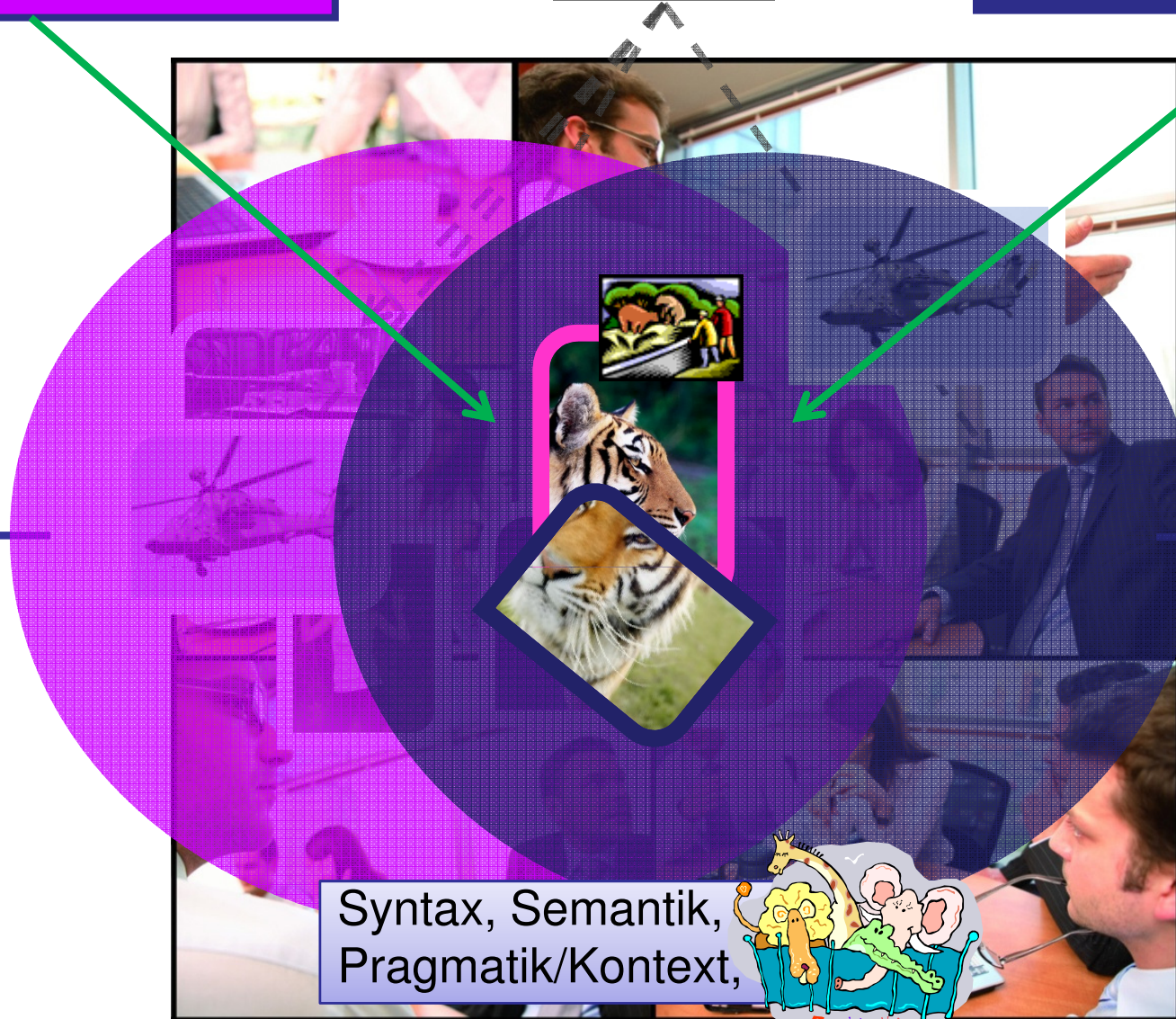
»People can't share knowledge if they don't speak a common language«
[Thomas Davenport (1993)]



Sender (Enkodiert)

»Tiger«

Empfänger (Dekodiert)



»People can't share knowledge if they don't speak a **common language**«

Einheitliche Spezifizierung der Symbole
zum Referenzieren der
Begriffe (Konzepte) und
Beziehungen (Relationen)

»Gemeinsame Sprache«, z.B.



»Tiger«

■ Symbole (Lexeme), Syntax

- »Task Force«, »Mission«, »**Auftrag**«, »**Aufgabe**«, ...
- ~~»Feind«, »Feindflieger«, »Feindflugzeug«, »Feind in Flugzeug«, »Enemy aircraft«, ...~~

■ Konzepte / Begriffe

- TaskForce, Mission, **Task**

■ Relationen

- leaderOf, hasWingman, memberOf, participantOf, ...

Vorteile der »gemeinsamen Sprache«

- Integriert und vereinheitlicht
 - Verschiedene **Quellen**, unterschiedliche **Repräsentationen**, verschiedene **Granularitätsstufen**
- Verschiedene Sichten auf Wissen möglich
 - Berücksichtigung von **Benutzer, Kontext**

»An ontology is an **explicit, formal specification of a shared conceptualization.**«
[GRUBER:1993]

»An ontology is an **explicit, formal specification of a shared conceptualization.**«
[GRUBER:1993]

- ***explicit***

- Semantik aller Begriffe definiert und repräsentiert

- ***formal specification***

- Maschinell les- und interpretierbar

- ***shared conceptualization***

- Gemeinsames Verständnis einer Domäne
 - Relevante Konzepte, ihre genaue Bedeutung, ihre symbolische Repräsentation in der Ontologie

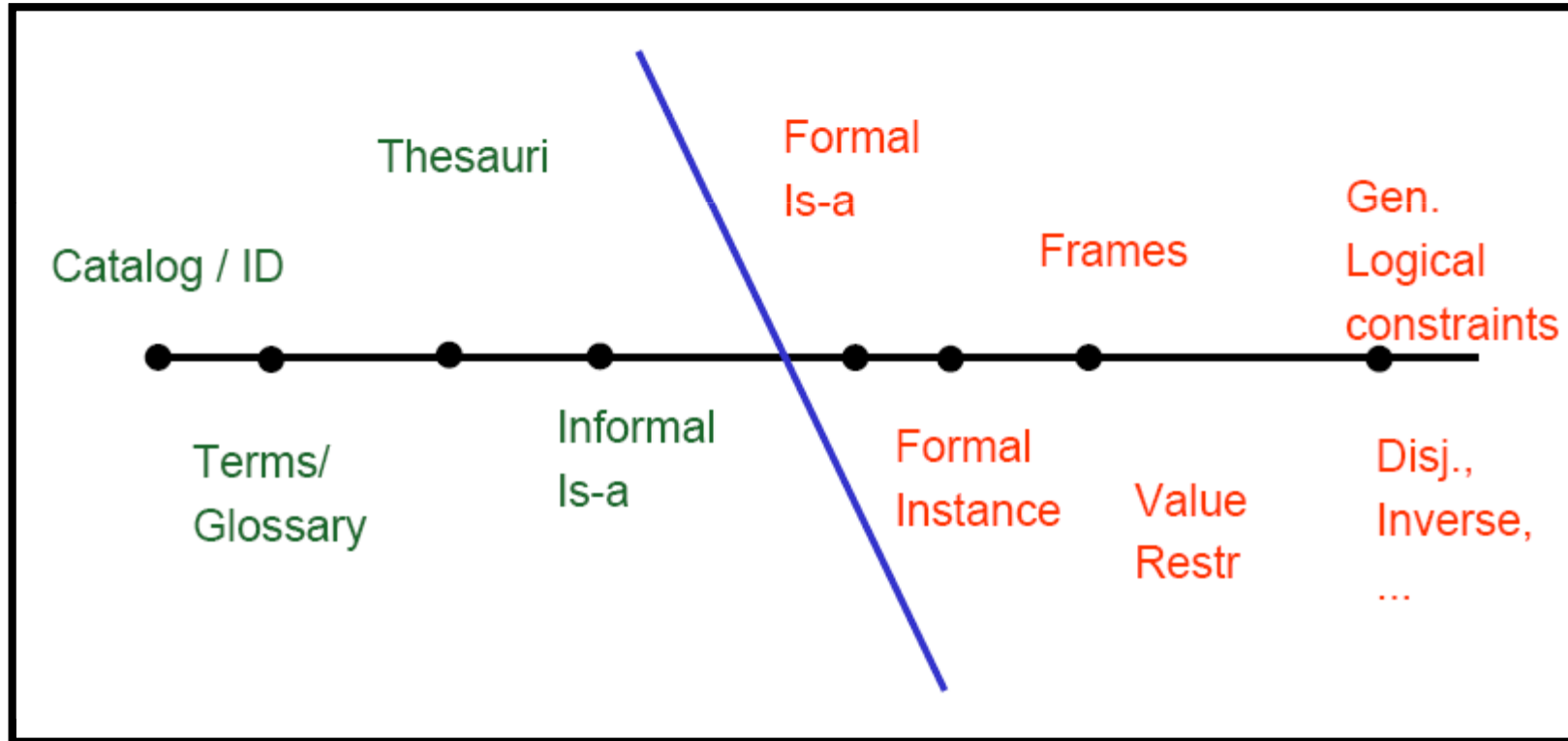
Einige formale Sprachen zur Repräsentation von Ontologien

- LOOM
- CyCL
- **F-Logic**
- Conceptual Graphs
- Ontolingua
- KIF
- **RDF**-Schema (W3C)
- UML: Unified Modeling Language
- DAML / OIL: Ontology Interchange and Inference
- **OWL**
- ...

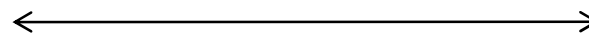
Beispiel (ein Teil der Ontolce Ontologie in F-Logic)

Konzepte	Beziehungen	Regeln
Object []. Person :: Object. Employee :: Person. Manager :: Employee. Consultant :: Employee. Project :: Object. Company :: Object. Manufacturer :: Company. FinanceComp :: Company. Insurer :: FinanceComp. LifeInsurer :: Insurer. Bank :: FinanceCompany. Location :: Object.	Person [firstName =>> String; lastName =>> String; email =>> String; phone =>> String; participantOf =>> Project; hasCompExperience =>> Company; address =>> Location] Project [projectname =>> String; projectgoal =>> String; client =>> Company; member =>> Person; leader =>> Person].	FORALL Proj1, Pers1 Proj1 : Project [member ->> Pers1] ↔ Pers1 : Person [participantOf ->> Proj1]. FORALL Pers1, Proj1, Comp1 Proj1 : Project [member ->> Pers1, client ->> Comp1] → Pers1 : Person [hasCompExperience ->> Comp1].

Ausdrucksstärke, Typen / Kategorien

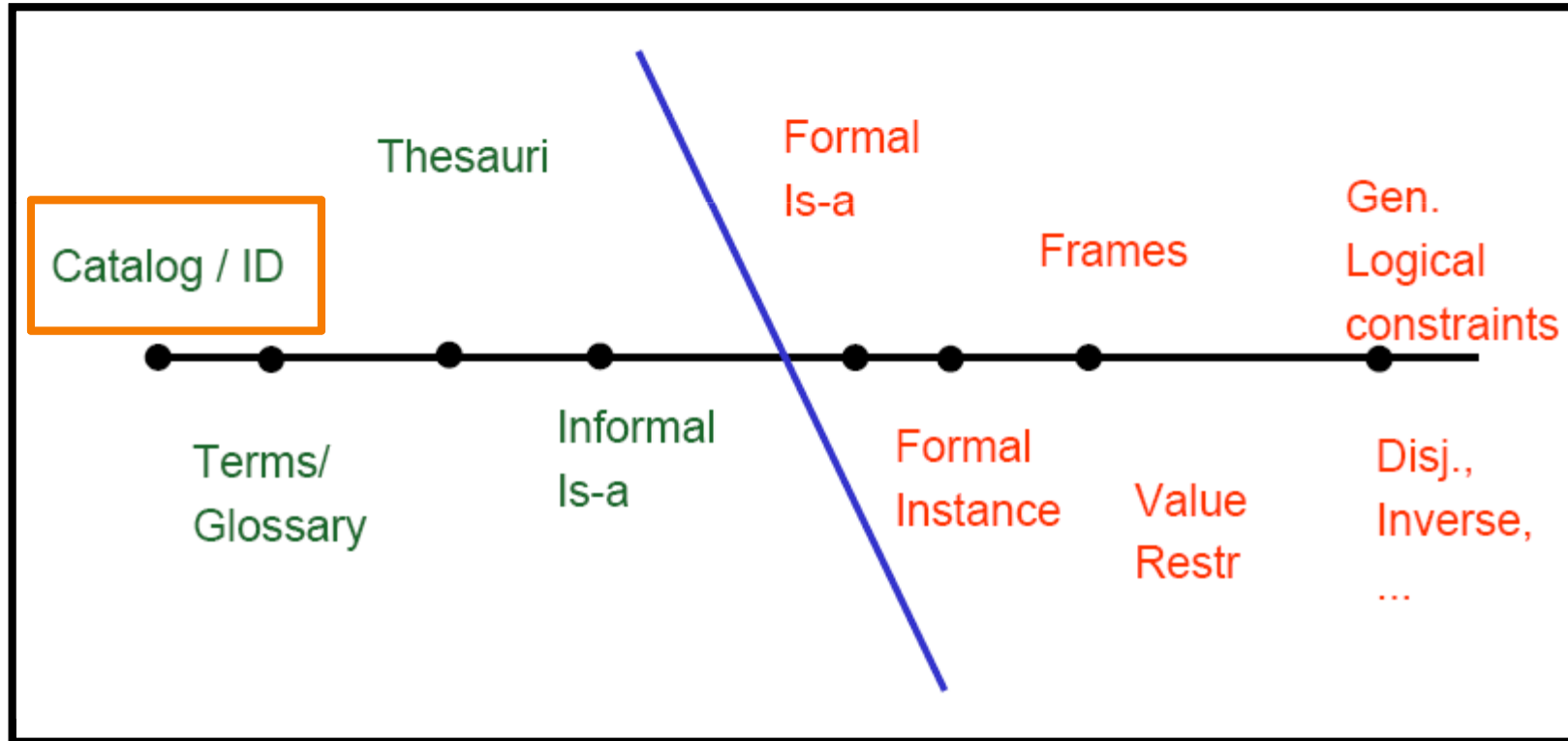


Lightweight ontologies



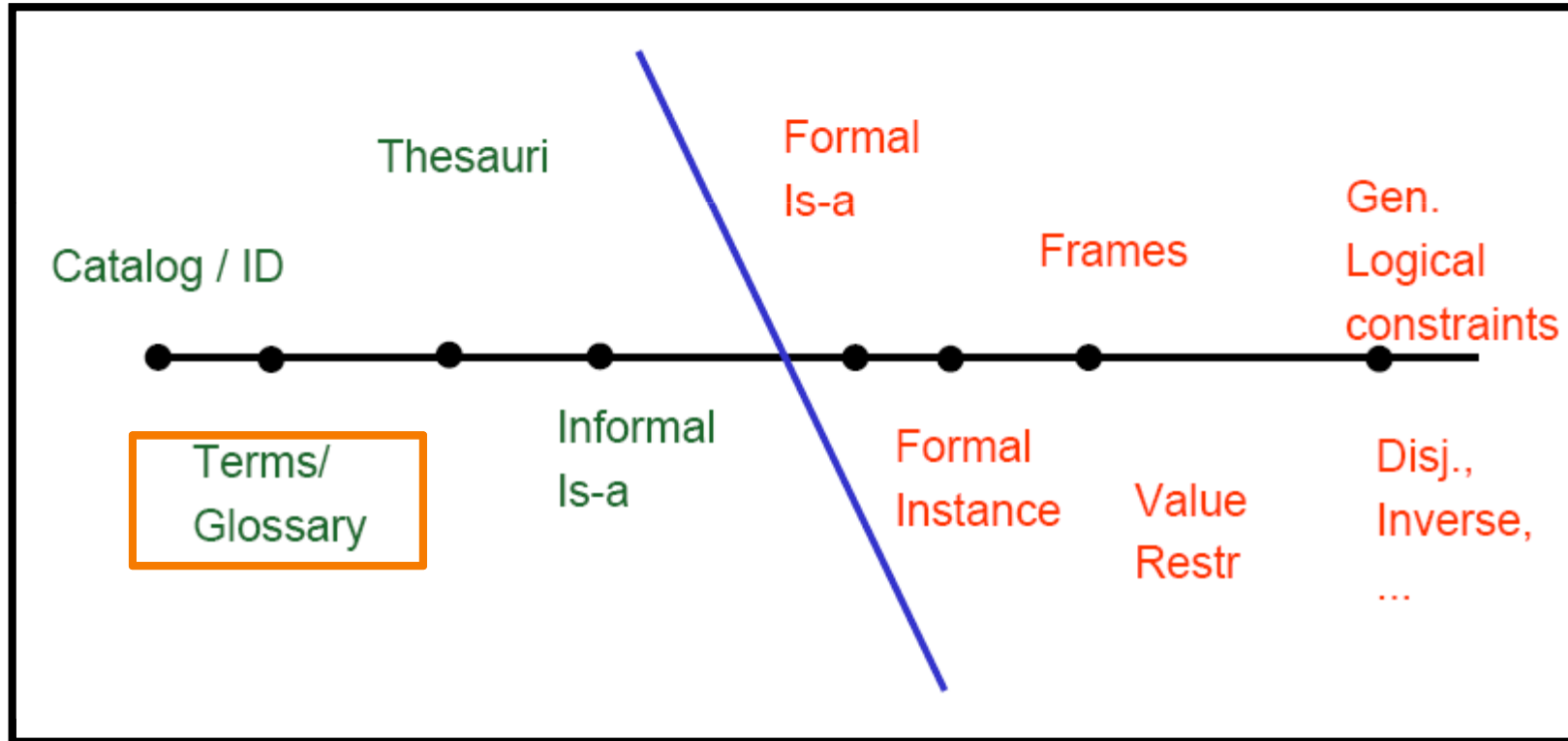
Heavyweight ontologies

Ausdrucksstärke, Typen von Ontologien



Katalog/Kontrolliertes Vokabular: Endliche Menge von Begriffen

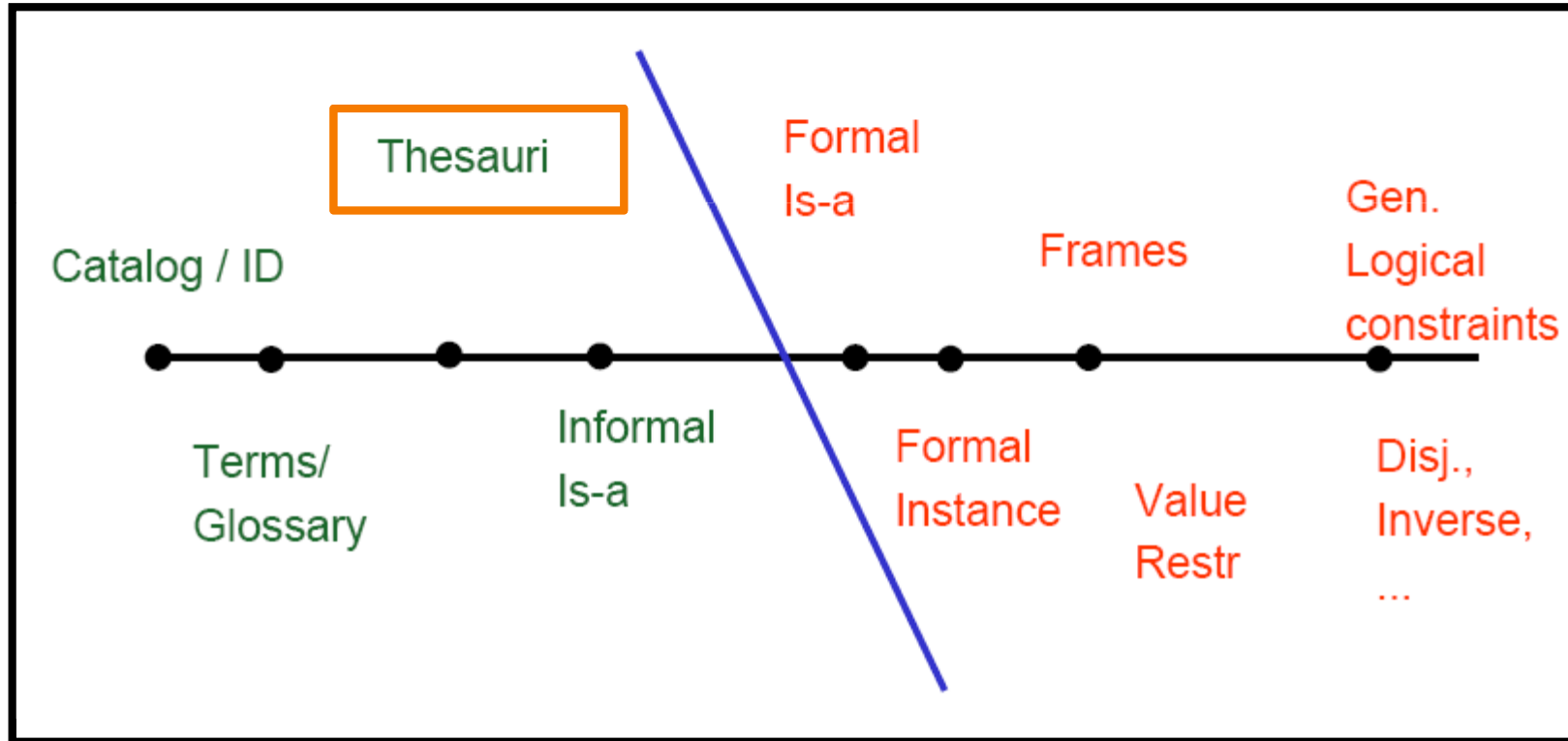
Ausdrucksstärke, Typen von Ontologien



Glossar:

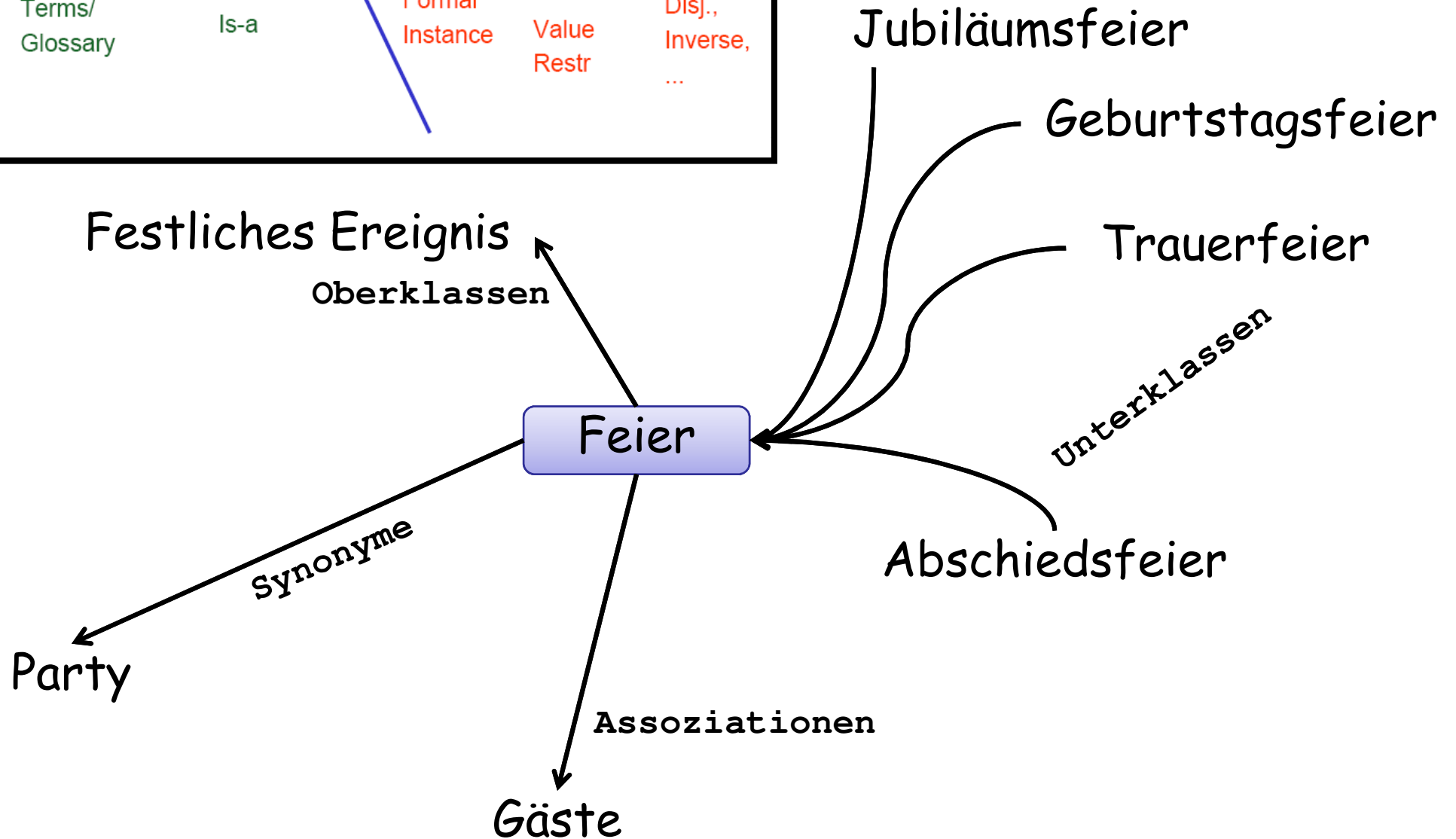
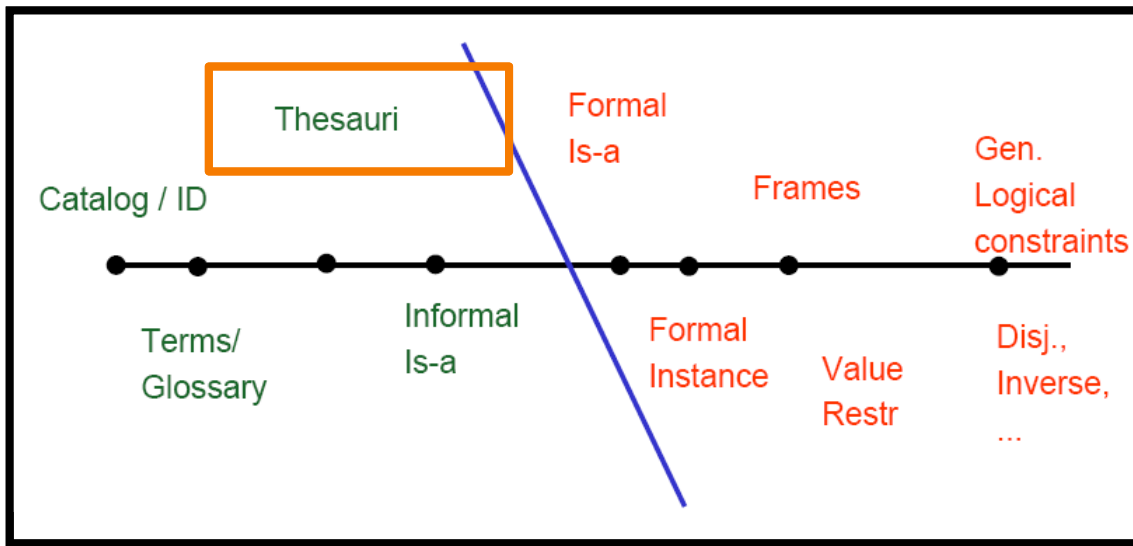
Endliche Menge Begriffe mit Bedeutungsbeschreibung. Informelle Beschreibung in natürlicher Sprache

Ausdrucksstärke, Typen von Ontologien

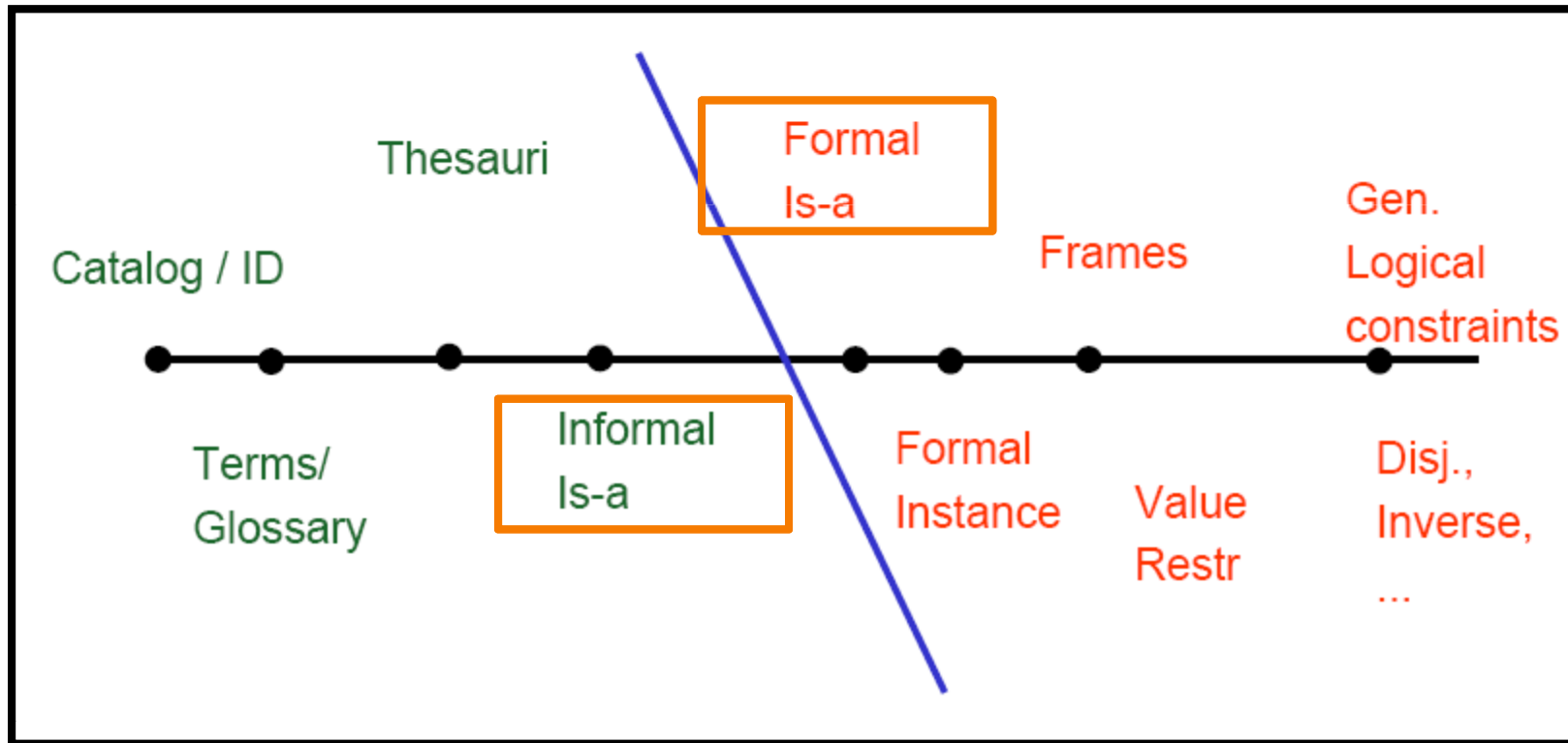


Thesaurus:

- Kontrolliertes Vokabular.
- Vernetzung der Begriffe über Relationen. Gängige Relationstypen:
Synonymie, Homographie, Antonymie, Taxonomie, Assoziation



Ausdrucksstärke, Typen von Ontologien



Informelle Taxonomie (informell): Explizite Hierarchie von Klassen, Unterklassenbeziehung nicht strikt eingehalten (Z.B. Yahoo)

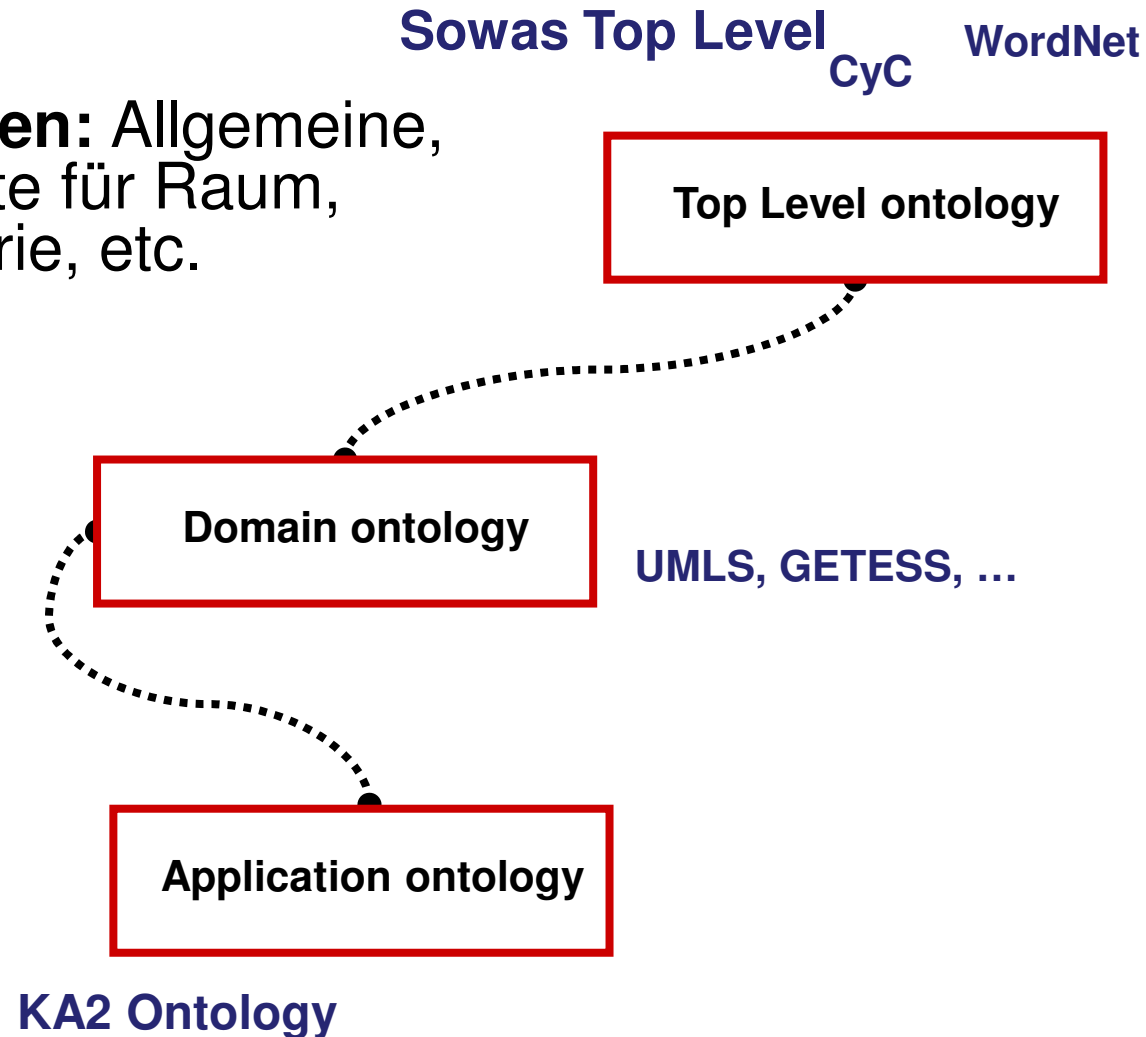
Formale Taxonomie: Explizite Hierarchie von Klassen, Unterklassenbeziehung werden strikt eingehalten wird (Transitivität / Instanz). Ggfs. auch neben der Unterklassenbeziehung: instance-of Vernetzung.

WordNet

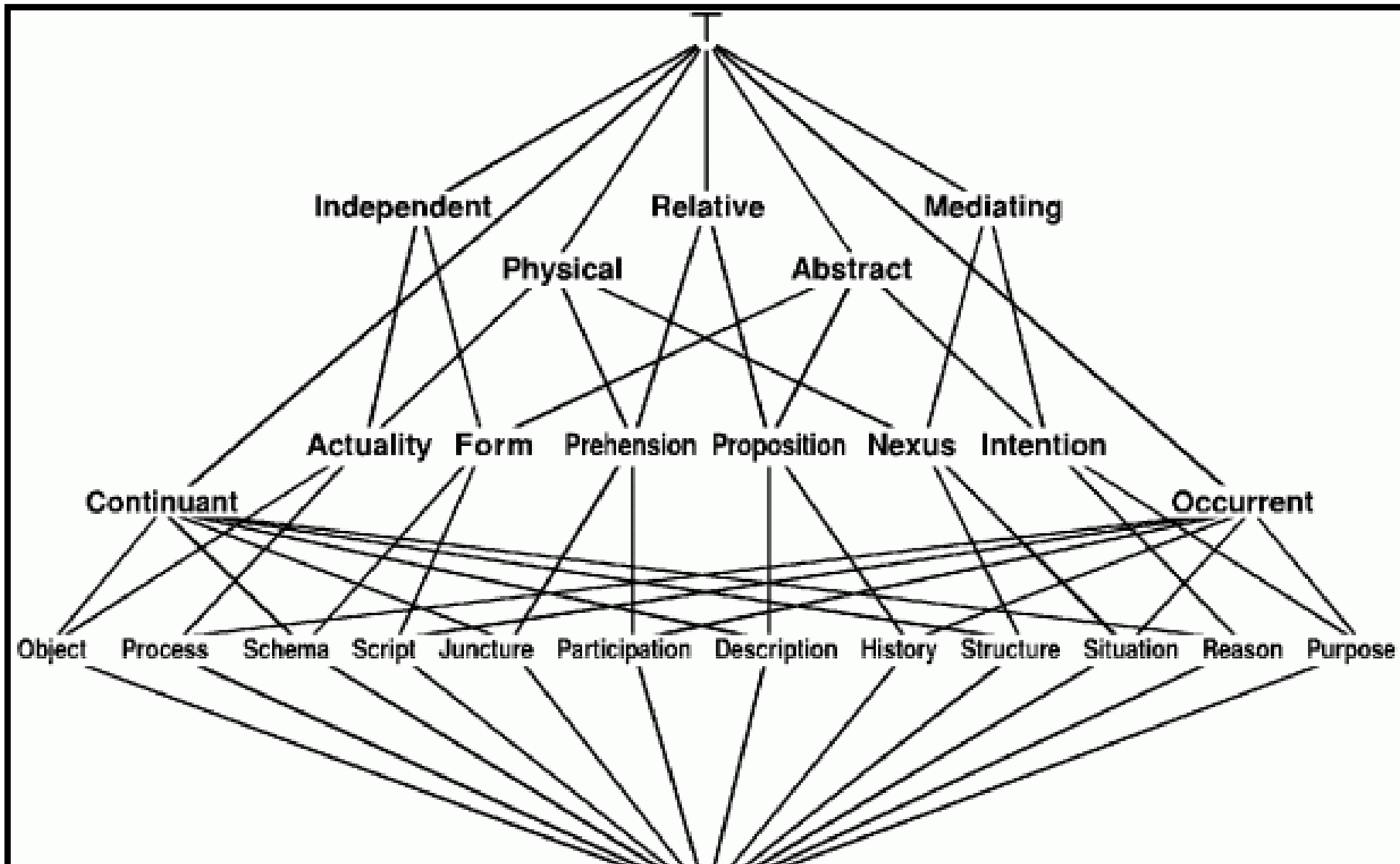
- Wörterbuch mit semantischen Relationen zwischen den Wortbedeutungen
- Organisation der Bedeutungen: »Synsets«, und diese wiederum nach Wortart: Nomen (N), Verben (V), Adjektive (Adj), Adverbien (Av)

Ontologie-Typen: Top-level, Domain, Application

Top-level Ontologien: Allgemeine, generische Konzepte für Raum, Zeit, Ereignis, Materie, etc.

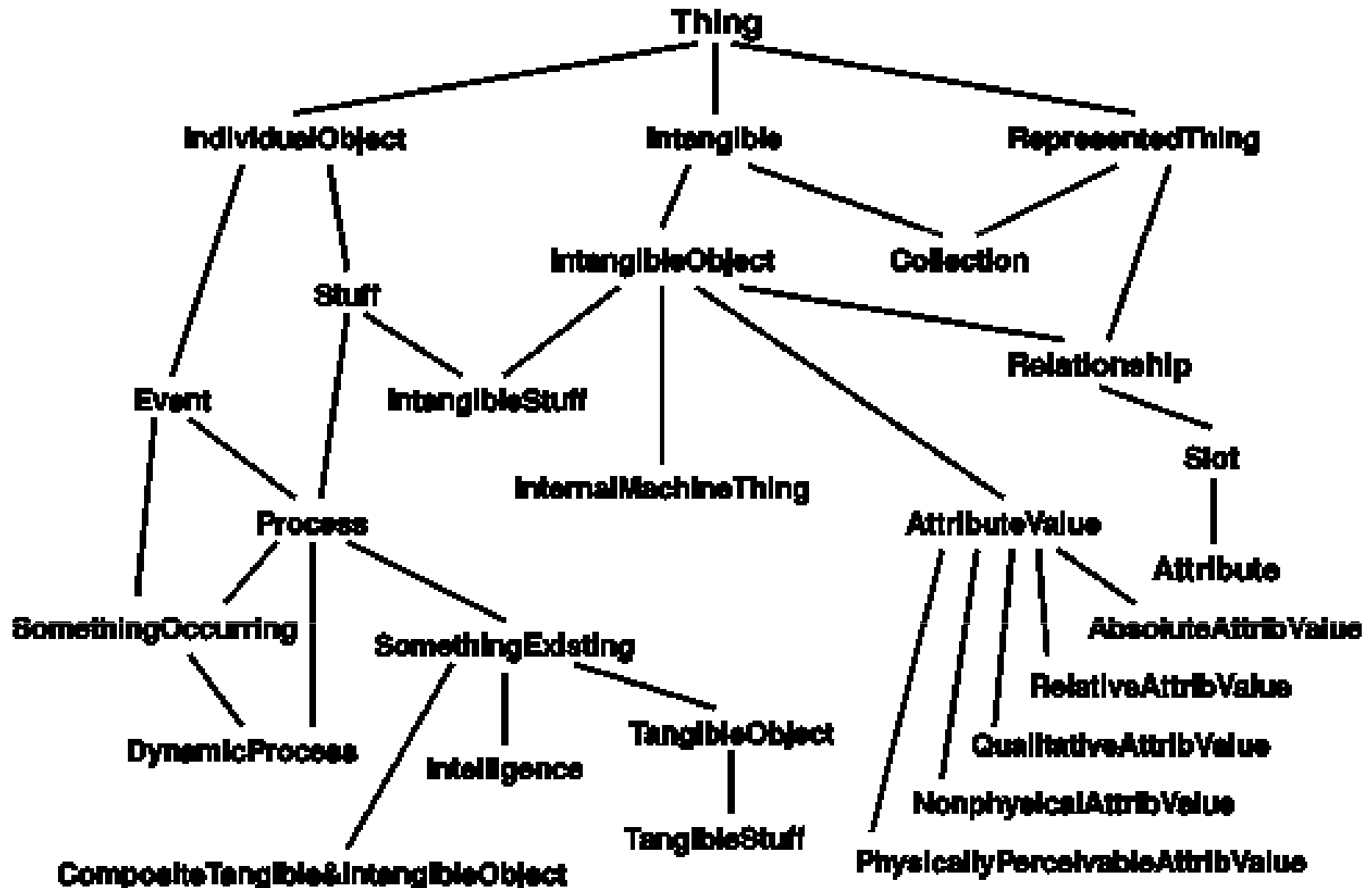


Sowas Top-level Ontologie



John F. Sowa's Top-level Ontologie mit 27 Konzepte als Verband (engl. *lattice*)

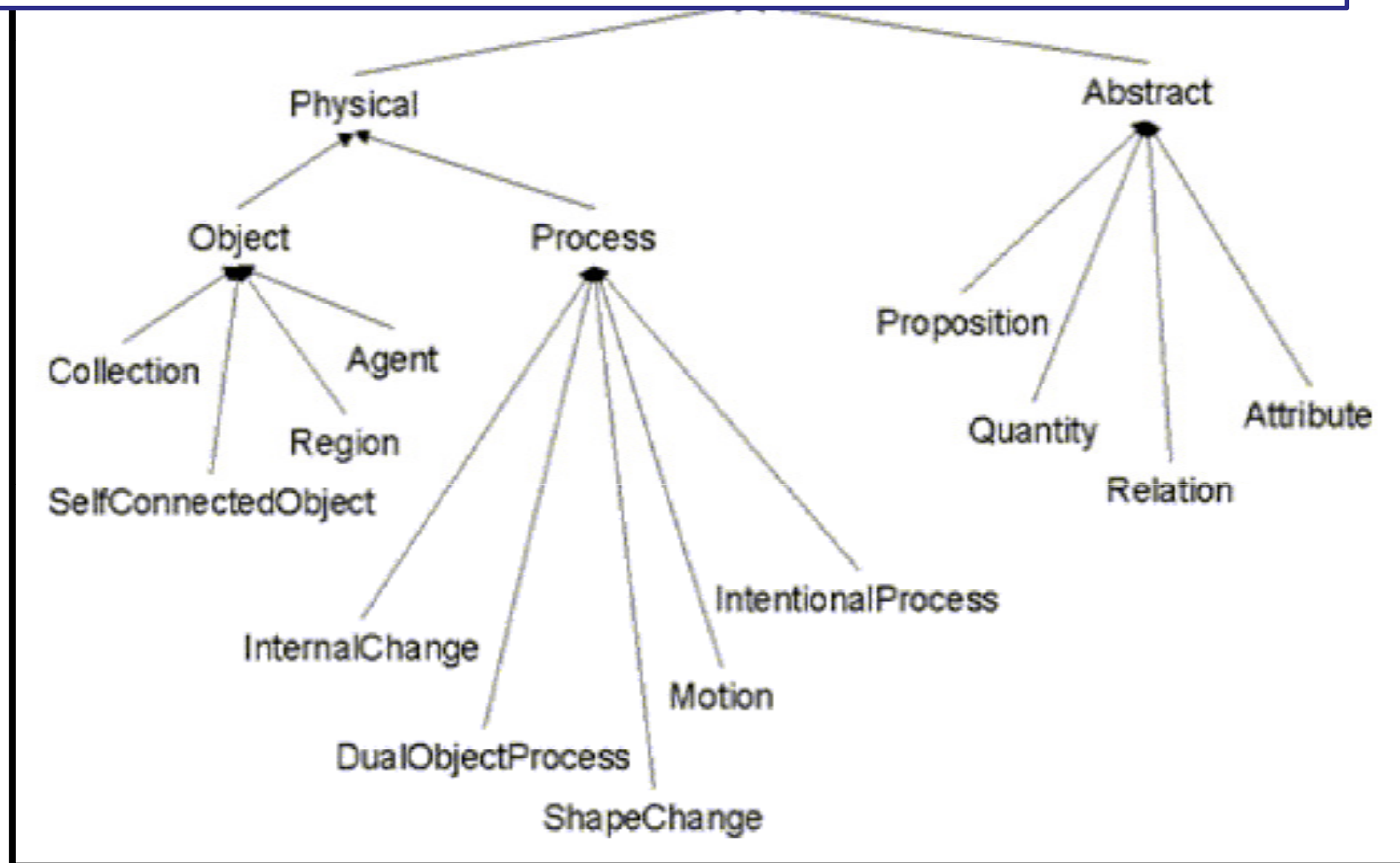
Cycs Top-level Ontologie



Top-level Ontologie: Standard Upper Merged Ontology (IEEE SUO Working Group)

An [...] is limited to concepts that are [...] generic [...], and therefore are general enough to address [...] a broad range of domain areas.

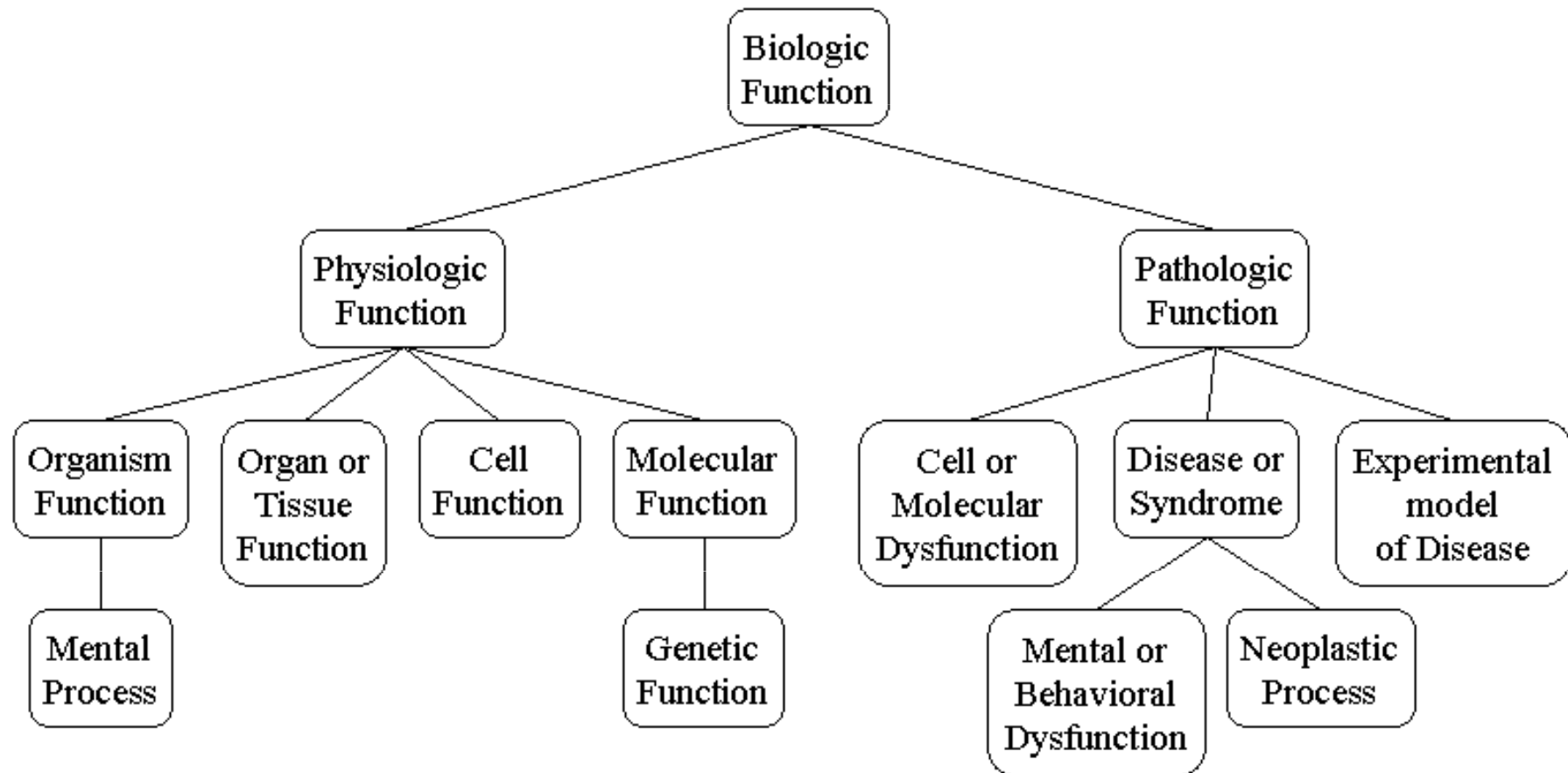
<http://suo.ieee.org/>



Domain ontology: **UMLS (Unified Medical Language System)**

- National Library of Medicine (NLM)
- Datenbank mit med. Terminologie
- Begriffe verschiedener med. Datenbanken (MEDLINE, SNOMED, etc.) vereint, so dass verschiedene Begriffe dieselben Konzepte referenzieren.
- <http://www.nlm.nih.gov/research/umls/>

Domain ontology: UMLS Ontologie (Ausschnitt)



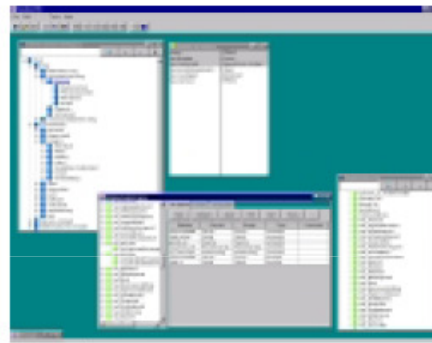
Ontologien für NLP

- Disambiguierung
- Koreferenz-Resolution (»Wer ist mit »er«/»sie«/»ihm« im Text gemeint?«)
- Metonyme Resolution (Was ist Teil von was/wem?)
- GETESS (German Text Exploitation and Search System)
 - Suchmaschine, Ontologie
 - Abgeschlossen, leider geht die Demo nicht mehr: <http://www.getess.de/>

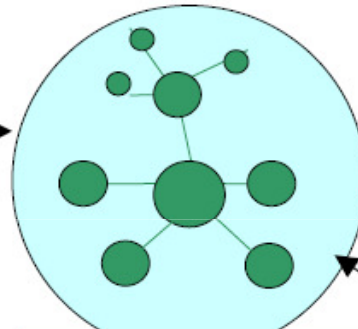
German Text Exploitation and Search System

[Staab et al., 1999]

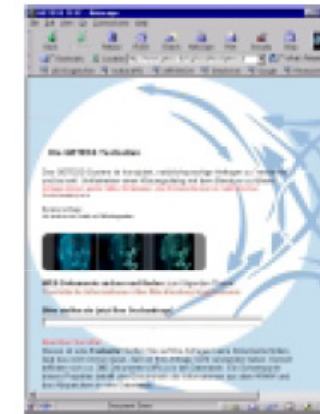
Ontology
Engineering



GETESS
Tourism
Domain
Ontology



Inference Engine



Information
Extraction

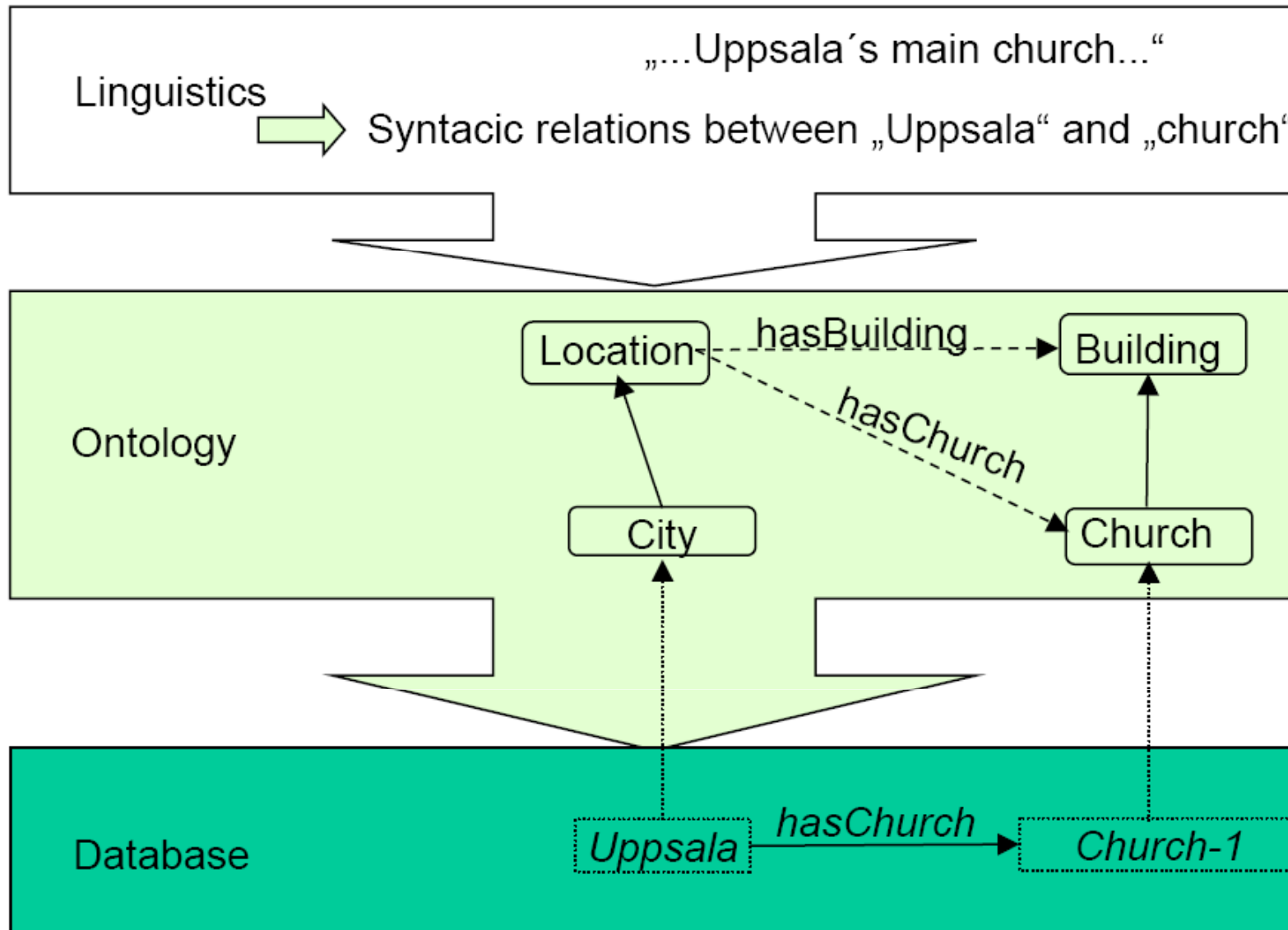
Database, Crawler,
Indexer, IR component



Web documents

[MAEDCHE:2000]

Translation into meaningful semantic relations



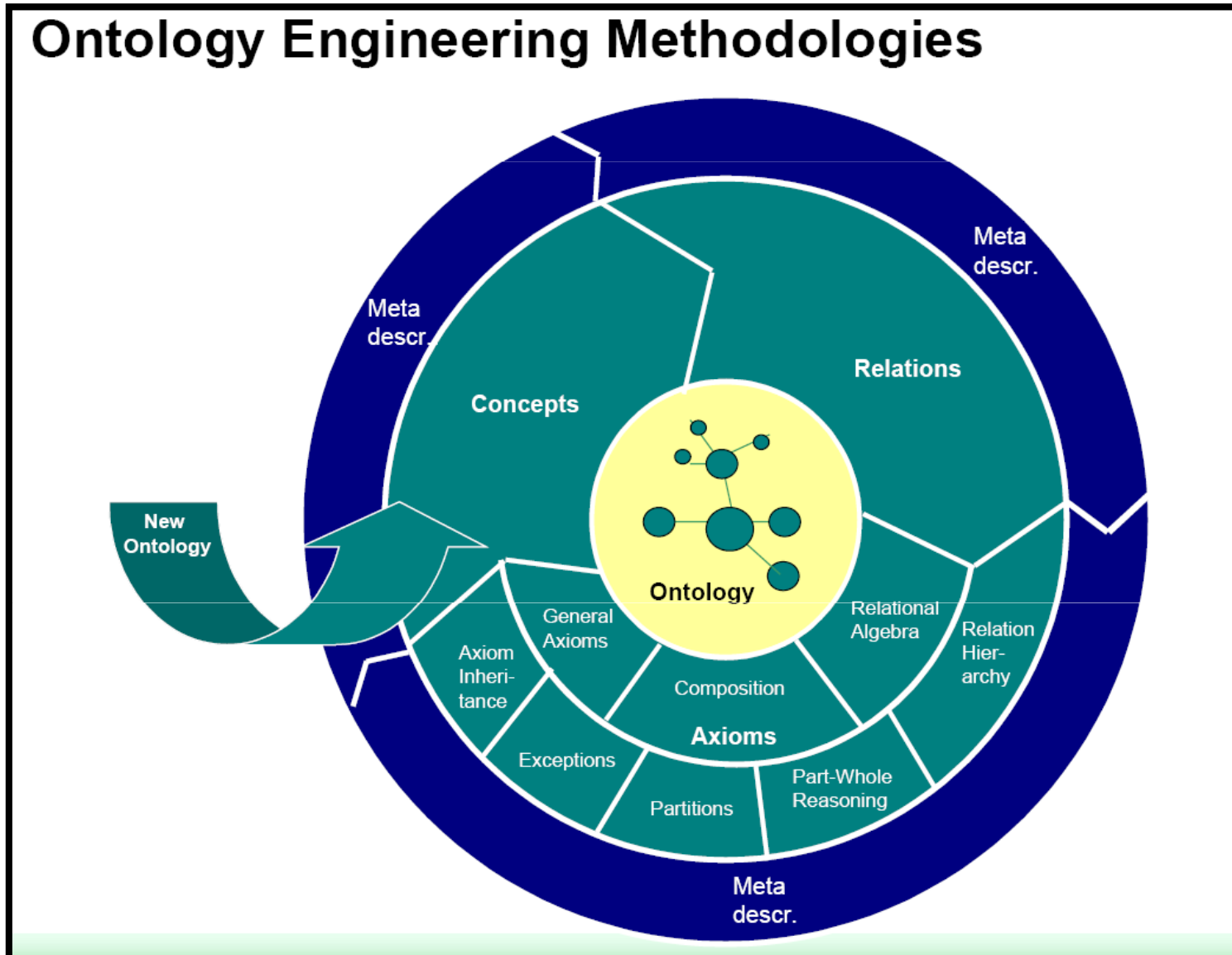
Inferenz-Maschinen

- Auch bei Ontologien:
 - Trennung von Wissensbasis und Verarbeitung
- Nicht-kommerzielle Inferenzmaschinen für Semantic Web Sprachen (RDF, OWL, ...):
 - [RACER](#)
 - [FaCT](#)
 - [FaCT++](#)
 - [Pellet](#)

Ontology Engineering

- (1) Zweck und Verwendung(-szenarien) der Ontologie identifizieren
- (2) Wissen erfassen, konzeptualisieren, formalisieren
- (3) Ontologie umsetzen in Rep.sprache
- (4) Existierende Ontologien integrieren
- (5) Evaluation der Ontologie
- (6) Dokumentation der Ontologie
- (7) Guidelines für den Aufbau, die Modifikation und Wartung der Ontologie

Ontology Engineering



Zweck identifizieren:

- (1) Warum wird die Ontologie benötigt?
 - (2) Was sind die möglichen Nutzungsformen?
- Beantwortung der Fragen haben die zentrale Funktion, den Rahmen für das zu modellierende Wissen zu stecken.
 - Grundlage, um bei der Generierung der Ontologie Hilfestellung zu folgenden Fragen zu leisten:
 - Ist der Begriff von Relevanz für die Ontologie?
 - Wenn ja, welche Eigenschaften des Begriffs müssen repräsentiert werden?

Wissen erfassen: *Top-down*, *middle-out*, *bottom-up* Strategien

- Das *top-down* Verfahren sieht vor, zunächst allgemeine Begriffe zu definieren, um sich davon ausgehend sukzessive zu den speziellen Konzepten vorzuarbeiten.
- Die *bottom-up* Methode verhält sich umgekehrt– erst das Spezielle, dann das Allgemeine.
- Die *middle-out* Vorgehensweise dagegen suggeriert, vorerst die fundamentalen Begriffe zu spezifizieren, und sukzessive konzentrisch die allgemeineren und spezielleren Konzepte zu ermitteln.

Wissen erfassen: Experten

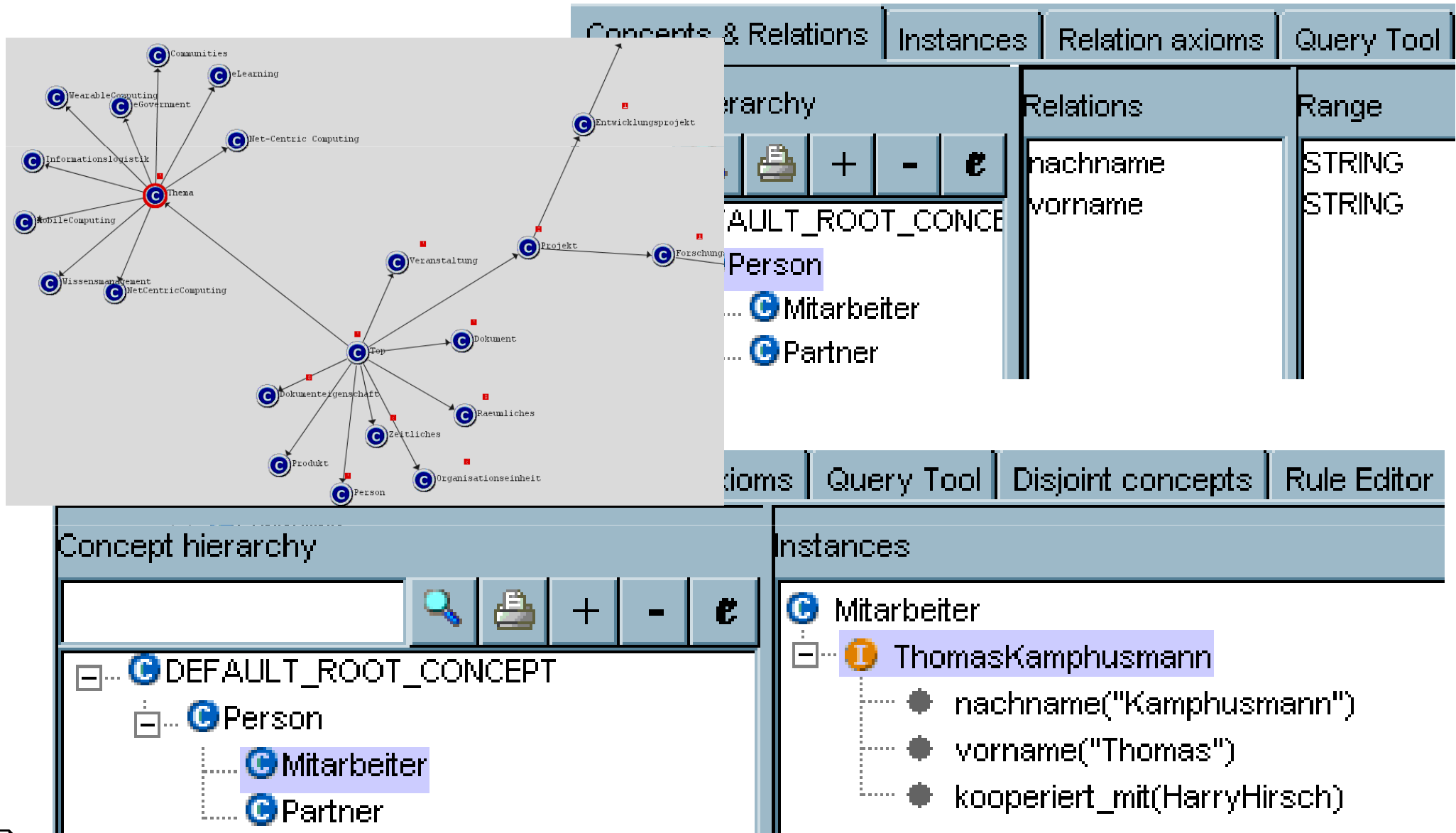
- Roter Faden für das zu erfassende Wissen: Festlegung durch den anvisierten Weltausschnitt, den Zweck und die involvierten Personengruppen der Ontologie.
- Wissen ist meistens heterogen verteilt auf Dokumente, Datenbanken und natürlich den involvierten Personen.
- Populärer Ansatz zur Detektion und Selektion von relevantem Wissen: Experten
 - Brainstorming der Experten
 - Interviews, Kompetenzfragen, freie Diskussion

Brainstorming

(IT-Ontologie mit Suche als haupts. Nutzung)

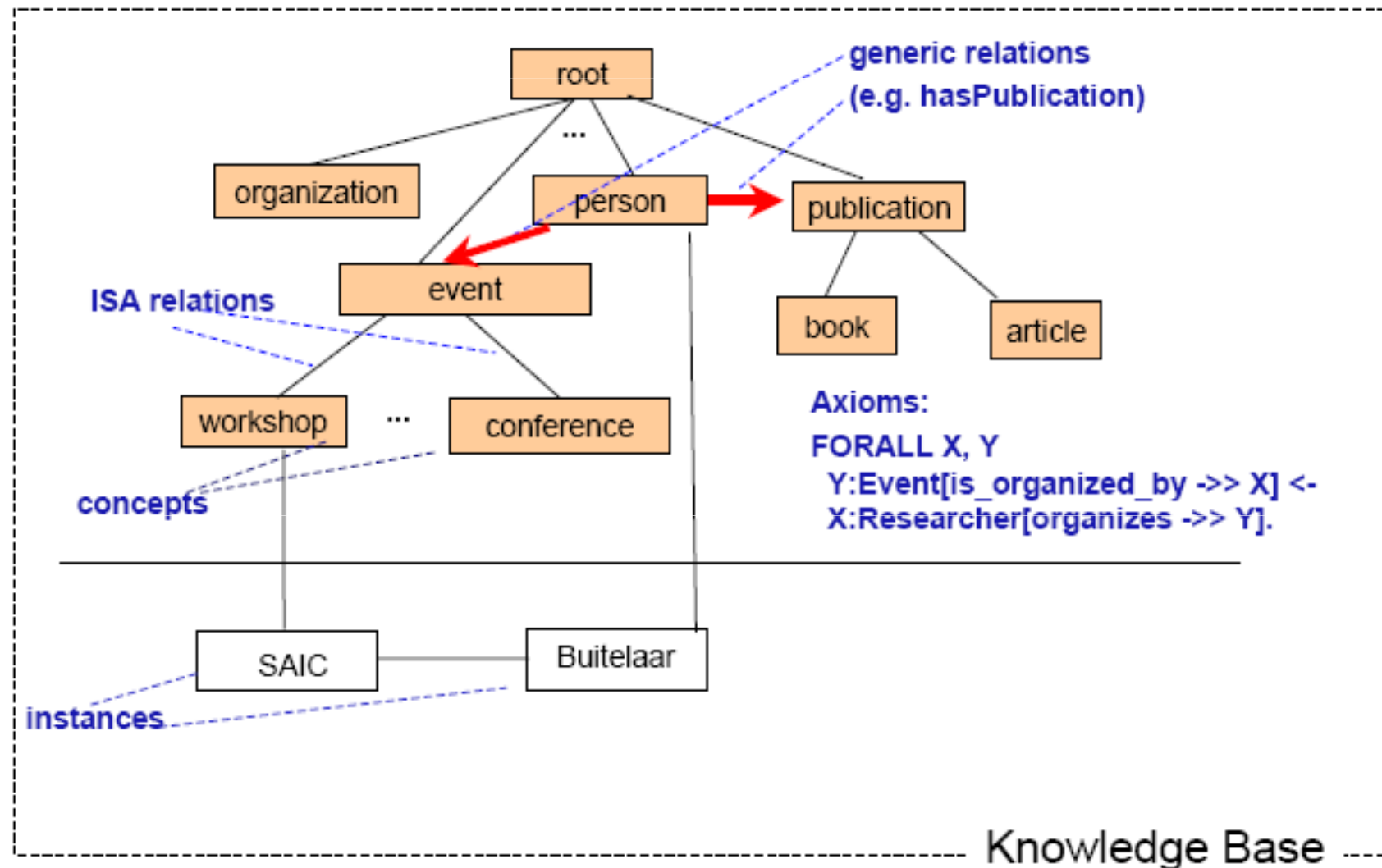
Abteilung, Person, Dokument, Buch, Artikel, Zeitschrift, Projektplan, Vertrag, Patent, Auftrag, Angebot, Lizenz, Meilenstein, Präsentation, Vorlage, Projekt, Forschungsprojekt, Entwicklungsprojekt, Tagung, Besprechung, Konferenz, Vortrag, Seminar, Hardware, Software, Netzwerk, Datenträger, Datenquelle, Raum, Etage, Gebäude, Adresse, Thema, Wissensmanagement, Informationslogistik, Organisation, eGovernment, eLearning, Rechner, Drucker, PDA, Telefon, Bildschirm, Notebook, Modem, Scanner, Workshop, Applikation, Systemsoftware, Betriebssystem, Browser, Entwicklungsumgebung, Multimedia, Groupware, Buchhaltung, EDV, Webserver, DBMS, DMS, Kommunikationstools, Messenger, VoIP, Windows, UNIX, Linux, Dokumentformat, Dokumenttyp, Mitarbeiter, Student, Auszubildender, Kunde, Partner, Webseite, Auftraggeber, Ansprechpartner, Produkt, Zeitpunkt, Besprechung, Meeting, Zeitraum, Zeitintervall, Zeitpunkt, Session, Seminarraum, Wissenschaftler, Akademiker, PhD Student, Mobiltelefon, Java, C++, Internet, Intranet, Beratung, IT, Informatik, Ontologie, WAN, LAN, Textverarbeitung, Tabellenkalkulation, OCR, Papier, Extranet, Kundengespräch, Mitarbeitergespräch, Kundenakquise, Reise, ...

Modellierung mit Hilfe von Werkzeugen (OntoEdit)



Domain ontology: Spezifische Axiome

Domain specific ontology = Object Model + Axioms



Axiom: Beispiel

- Wenn eine Person X ein Projekt Y leitet, dann besitzt Person X auch *Know-How* über die Themen Z, die das Projekt Y behandelt.
 - Herr Uwenknecht ist eine Person und leitet das Projekt Ontoware.
 - Ontoware behandelt die Themen Ontologien und Wissensportale.
 - ➔ Herr Uwenknecht besitzt *know-how* über die Themen Ontologie, Wissensportale und Wissensmanagement.

Axiom: Beispiel

- Wenn X mit Y kooperiert, dann kooperiert Y auch mit X.
 - Harry kooperiert mit Gilda.
 - Gilda kooperiert mit Harry.
- Wenn X eine Klage gegen Y gewinnt, dann verliert Y die Klage gegen X.
 - Porsche gewinnt die Klage gegen BMW.
 - BMW verliert die Klage gegen Porsche.

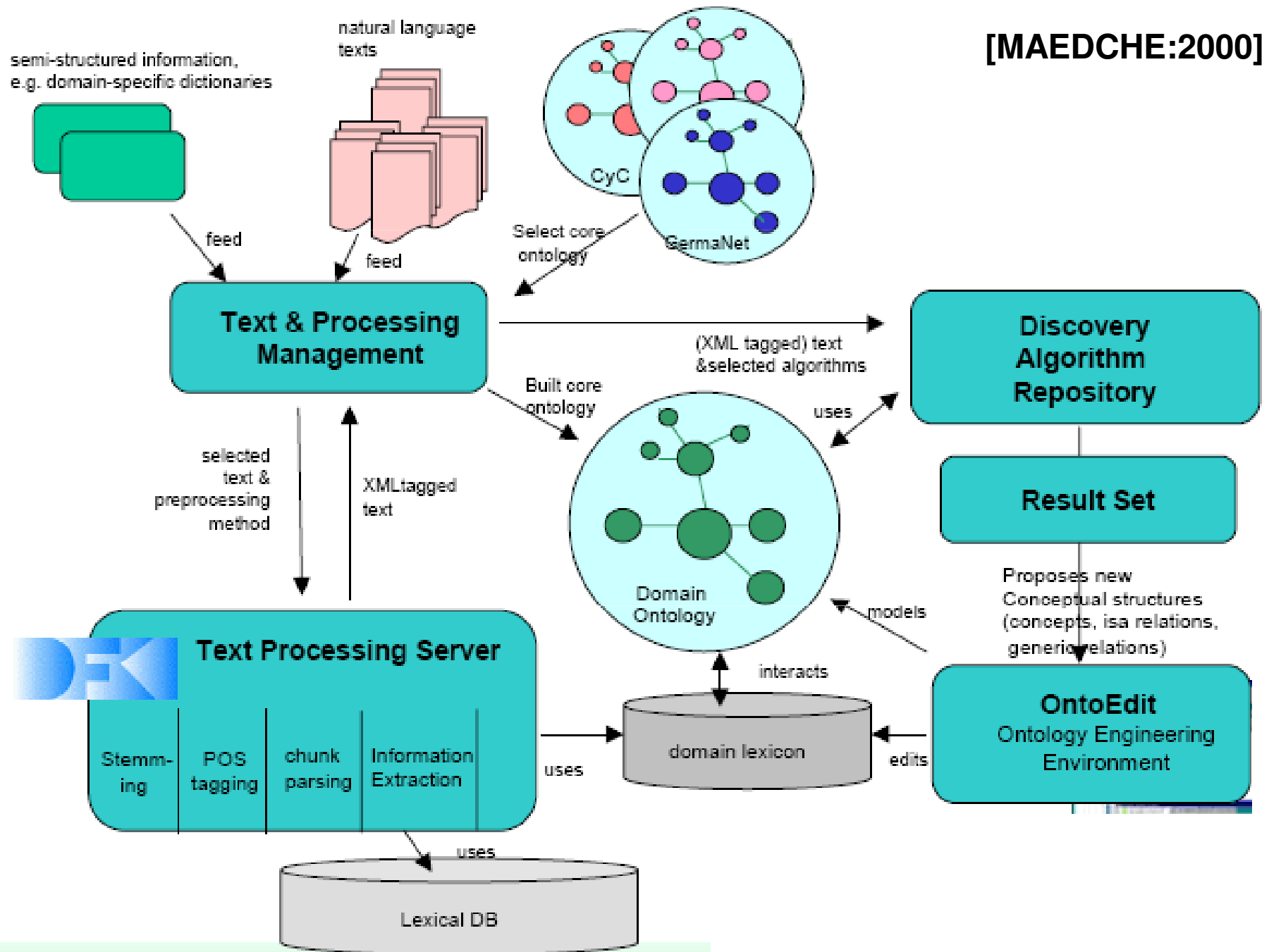
Axiome in F-Logic

```
FORALL X, Y
  ( X["http://www.do.isst.fhg.de"#arbeitet_an_Produkt->>Y] )
  <->
  ( Y["http://www.do.isst.fhg.de"#hat_Ansprechpartner->>X] ) .
```

```
FORALL X, Y
  ( X["http://www.do.isst.fhg.de"#hat_Referenzkunden->>Y] )
  <-> ( Y["http://www.do.isst.fhg.de"#bezieht_Produkt->>X] ) .
```

```
FORALL X, Y
  ( X["http://www.do.isst.fhg.de"#resultiert_aus_Projekt->>Y] )
  <-> ( Y["http://www.do.isst.fhg.de"#implementiert_Produkt-
  >>X] ) .
```

(Semi-)Automatisches Ontology Engineering



Werkzeuge/Entwicklungsumgebungen fürs Ontology Engineering

- <http://protege.stanford.edu/download/prerelease/full/>
- <http://www.mindswap.org/2004/SWOOP/>
- <http://owl-eclipse.projects.semwebcentral.org/>
 - **SWeDE** (Semantic Web Development Environment)
- <http://owlseditor.semwebcentral.org/>
- Kowari, RDFLib, Jena, Sesame
- (Kommerziell) Ontoprise (<http://www.ontoprise.de>)

Referenzen

[DAVIS:2003]

Dieter Fensel, Frank van Harmelen, Ian Horrocks. In John Davies et al. (ed.), ***Towards the Semantic Web: Ontology-Driven Knowledge Management***. Wiley, 2003.

[DIENG:2000]

Rose Dieng, Olivier Corby, editors. ***Knowledge Engineering and Knowledge Management: Methods, Models and Tools, Proc. of the 12th International Conference, EKAW'2000***, Juan-les-Pins, 2-6 octobre 2000. Springer-Verlag, LNAI n.1937.

[GRUBER:1993]

T. R. Gruber. ***Toward principles for the design of ontologies used for knowledge sharing***. Presented at the Padua workshop on Formal Ontology, March 1993, later published in *International Journal of Human-Computer Studies*, Vol. 43, Issues 4-5, November 1995, pp. 907-928.

[LASSILA: 2001]

Lassila, O., McGuinness, D.: ***The Role of Frame-Based Representation on the Semantic Web***, Technical Report KSL-01-02, Knowledge Systems Laboratory, Stanford University, California, USA, 2001.

[MAEDCHE:2000]

Alexander Maedche. Vortrag: ***Tutorial: Development and Applications of Ontologies***. Konferenz / Veranstaltung: FGML 2000, 16.09.2000, Bonn.