# Applying lexical knowledge to improve search quality for a German legal information database Master thesis final presentation

Laura Altamirano Sainz - May 4th, 2014

Software Engineering for Business Information Systems (sebis)
Department of Informatics
Technische Universität München, Germany

wwwmatthes.in.tum.de

# Administration matters

- **Time:**

  October 15th, 2014 to April 15th, 2015

- **Supervisor:**

  Prof. Dr. Florian Matthes

- **Advisor:**

  Bernhard Waltl

# Agenda

1. **Motivation**
2. **Research questions**
3. **Research method**
4. **Demonstration**
5. **Evaluation**
6. **Conclusions**

# Motivation

- Integration of lexical information in legal searches

**Related work:**

- Ontologies integrated in the foreground of the systems
  - Interaction between the users and the lexical knowledge
- Other areas integrate lexical knowledge for searches: Biology
- Lexical knowledge integrated in the background

of the systems

- Query expansion

# Research questions

How can search quality be improved by lexical information for a legal database?

What mechanisms and methods are common in legal databases?

Which search mechanisms and methods can be enhanced by lexical information and how?

How can a implementation for a support search mechanism integrated with lexical knowledge look like?

# Research method

*"Systems are built to help people work better. They cannot be built well without understanding how people work" (Holtzblatt & Beyer, 1997)*

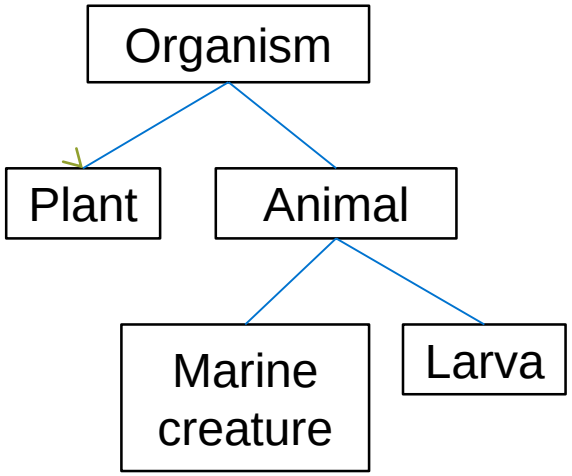
- **Needs assessment**
  - Interview with 6 experts in the legal domain
  - *Results:*
    - Lexical relations
      - Hyponyms
      - Troponyms
      - Siblings / related terms
      - Derivationally related terms

- **Search support mechanisms in legal databases analysis**
  - Comparison between 5 legal databases
  - *Search support mechanisms categories*
    - Query formulation / specification
    - Query reformulation
    - Integration of navigation of the results with search

## Search system

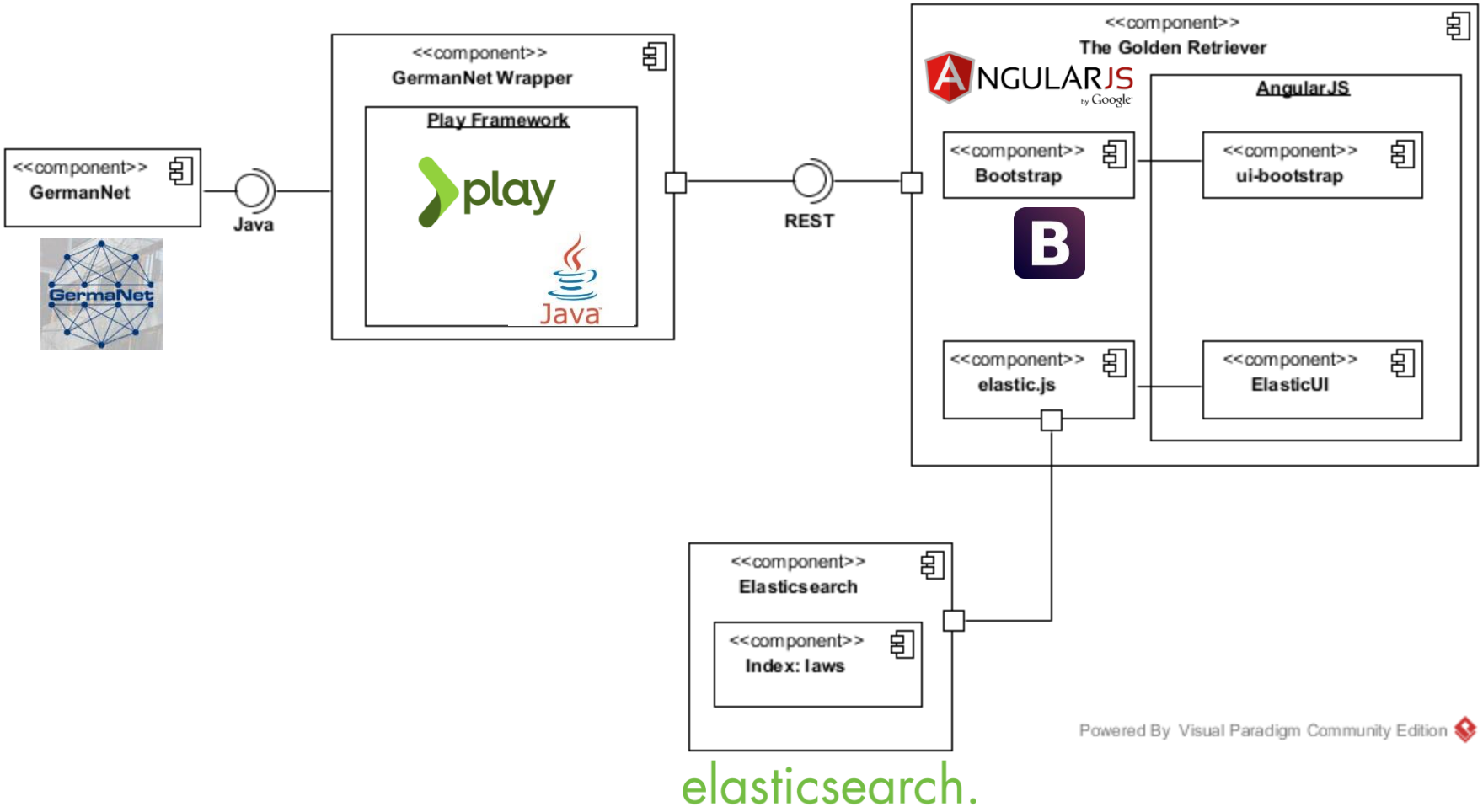
- Integration of the lexical knowledge as a search support mechanism in a search system
- *GermaNet*
  - Lexical database for German


- *Query expansion / refinement suggestions*
  - Query expansion with hypernyms
  - Query refinement with hyponyms



# Research method

## System architecture



# DEMONSTRATION

# Evaluation

- **Limitations**
  - Word sense disambiguation
  - Search context
  - Lexical database

- **Evaluation results – Expert interview**

| System advantages | Areas of improvement |
| --- | --- |
| • Integration of lexical information as a search support mechanism<br>• More than one lexical relation implemented<br>• Clean and clear interface | • Search context<br>• Personalization<br>• Explanatory mechanism for highlighted words |

# Conclusions

- Lexical information can improve searches
- Current search support mechanisms can be improved by lexical information
- Law practitioners show interest for this area
- Users are able to interact with the lexical information

***Outlook***

- Context
- Improving lexical database
- Personalization features

# Thank you for your attention!


**Laura Altamirano Sainz**

Technische Universität München
Department of Informatics
Chair of Software Engineering for
Business Information Systems

Boltzmannstraße 3
85748 Garching bei München

Tel +49.89.289.17124
Fax +49.89.289.17136

laura.altamirano@tum.de
wwwmatthes.in.tum.de

# References


**[1]** http://www.toddtransportation.com/countdown-to-a-successful-move-the-moving-checklist/

**[2]** http://www.horsesforsources.com/cognizant-051411

**[3]** AngularJS. https://angularjs.org/

**[4]** Elasticsearch. https://www.joyent.com/public-cloud/benchmarks/elasticsearch

**[5]** Play framework. https://www.playframework.com/

**[6]** Germanet. http://www.sfs.uni-tuebingen.de/GermaNet/

**[7]** Wordnet. http://wordnet.princeton.edu/wordnet/

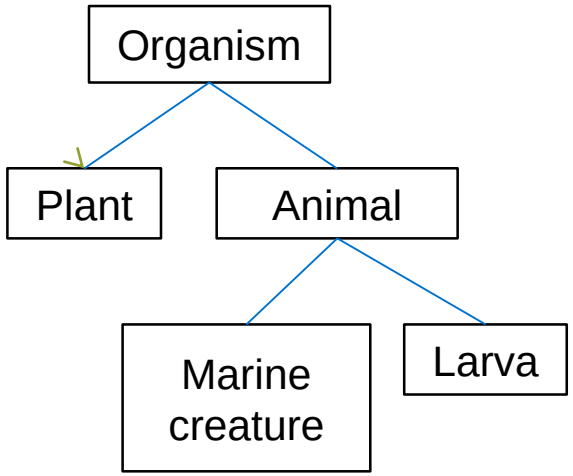
**[8]** Java logo. http://onsitepcsolution.com/wp-content/uploads/2014/08/java_tech.jpg

**[9]** Bootstrap logo. http://logonoid.com/bootstrap-logo/

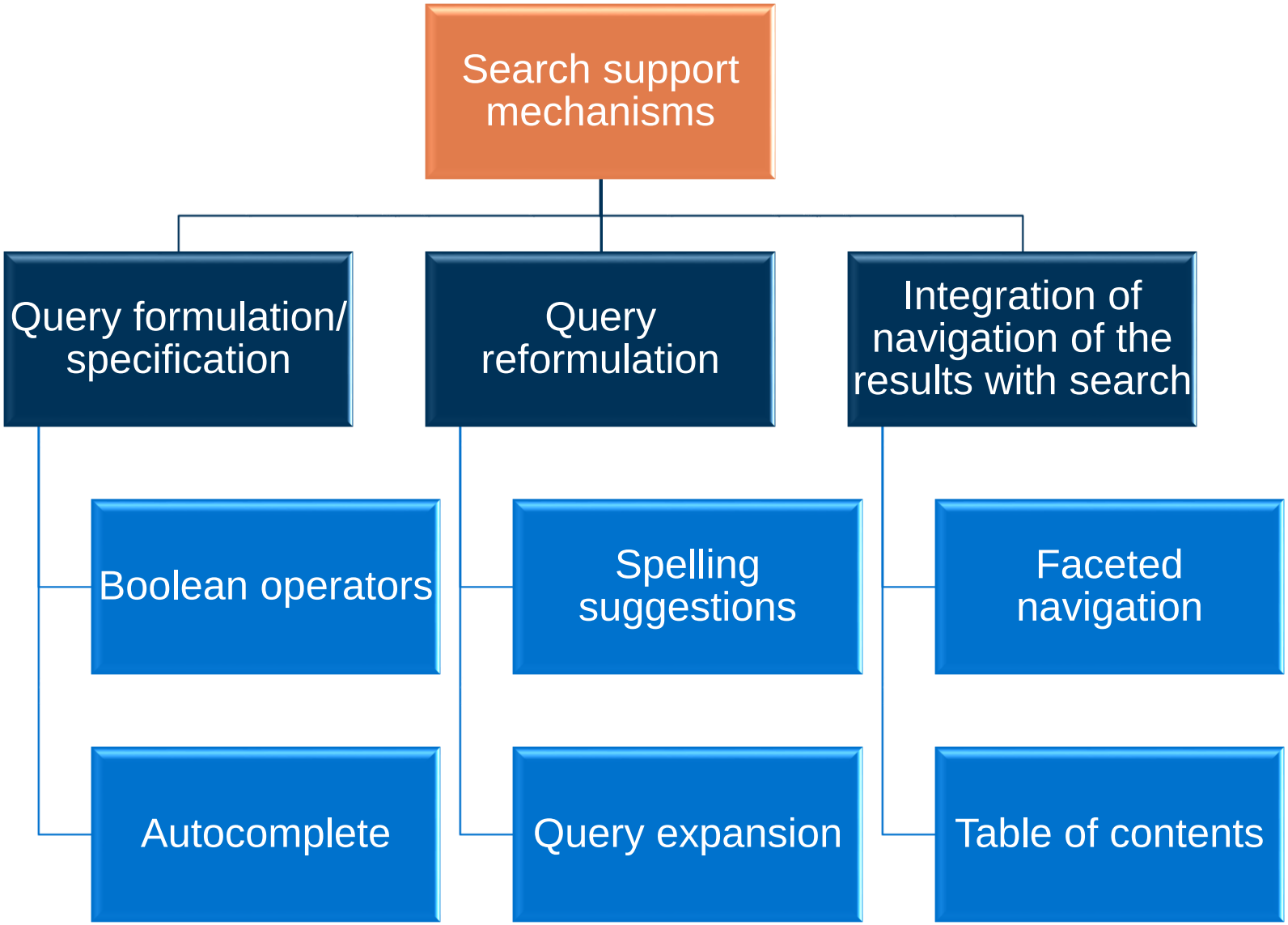
**[10]** http://www.many-roads.com/2014/12/13/mega-search-engine/

- **Lexical knowledge**
  - All what we know about a word
  - Relationships with other words
  - Ontological categories (Relations, hyponyms, hypernyms, synonyms,…)

  - **Lexical databases**
    1. **Wordnet**
       - 117 000 synsets
    2. **GermaNet**
       - 93 246 synsets



# Backup slides



## German legal information

- Number of German laws is increasing
- Frequently revised information
- Relevant information
  - To build up cases, for resolutions, etc…
- Style of writing legal documentation
  - Standard format

# Backup slides

- Main objectives
  - Finding out common search support mechanisms in legal databases
  - Assess the needs of the user
  - Integrate the lexical information as a search support mechanim in a system
  - The support tool must be intuitive to the user
  - Evaluate the system

# Backup slides

Search support mechanisms analysis in 5 legal databases

- JURION
- beck-online
- LexisNexis
- FindLaw
- LEXinform

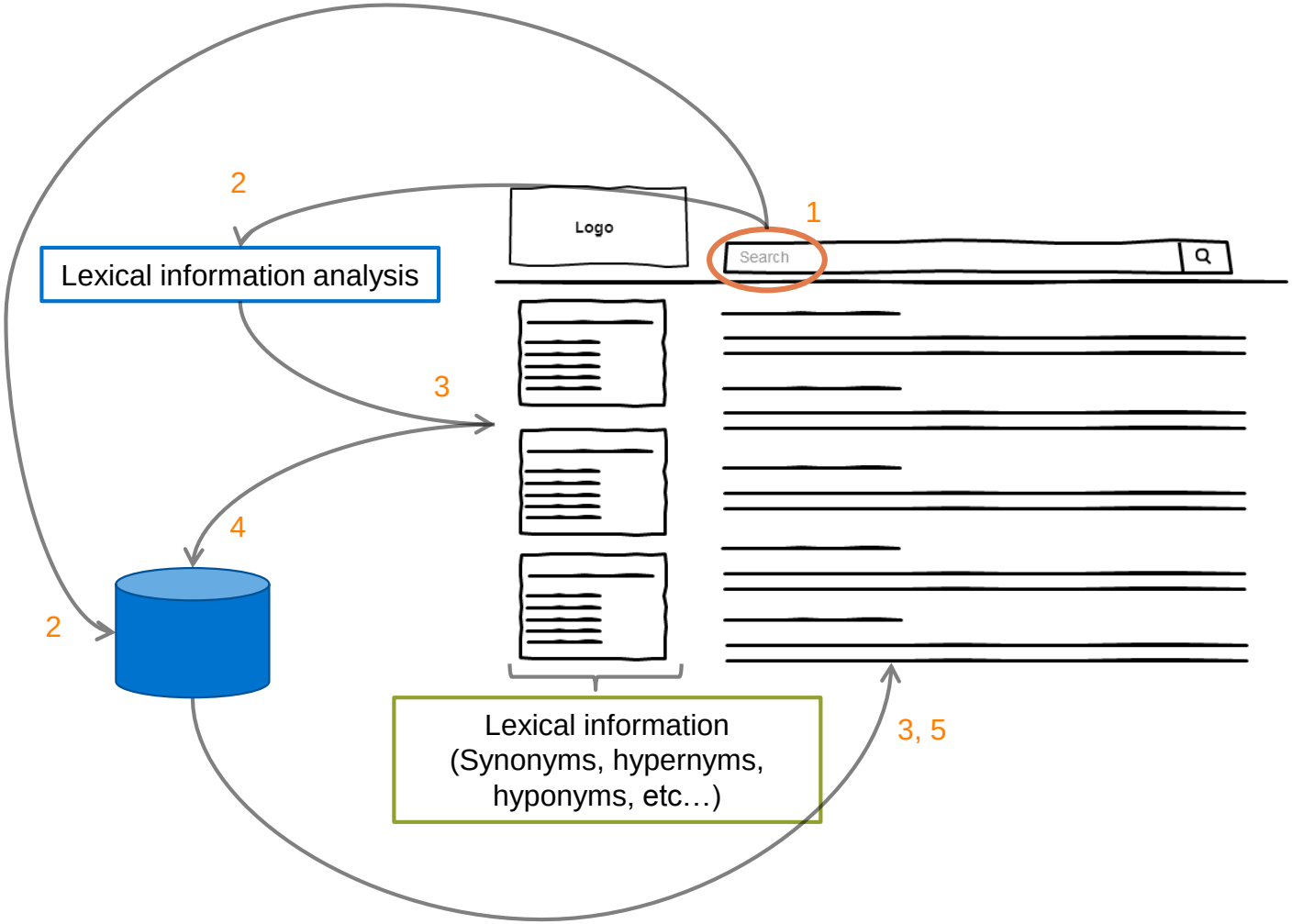
Select which ones are common

Select which ones can be enhanced by lexical information

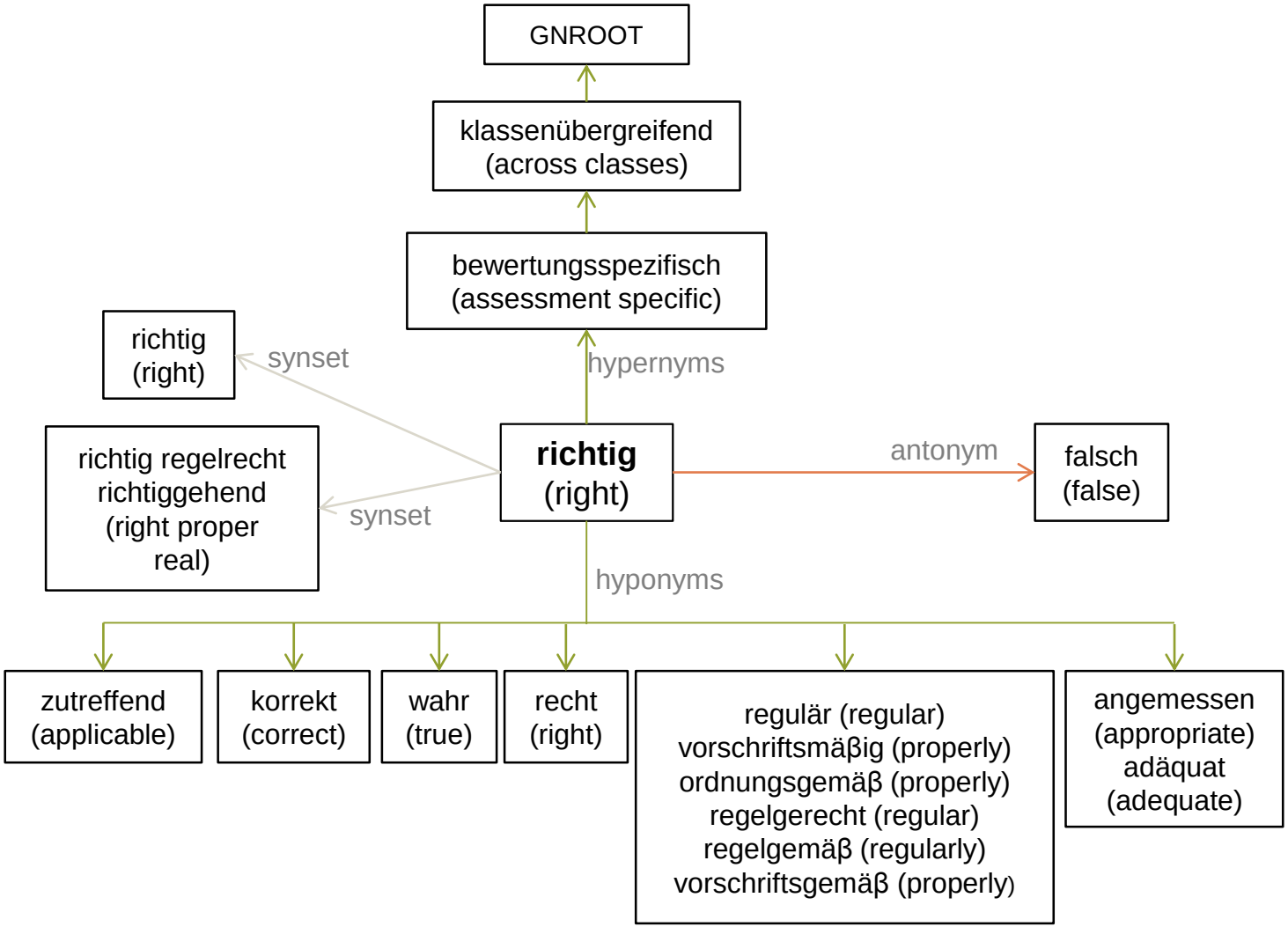
- For example:
  1. Autocomplete
  2. Faceted navigation

Needs assesment

## Process


2
Lexical information analysis
Logo
1
Search
3
4
2
3, 5
Lexical information
(Synonyms, hypernyms,
hyponyms, etc…)

## Lexical relations



## Sequence diagram