

*– A network of words*

Bolette Sandford Pedersen  
Center for Sprogteknologi





## DanNet: Joint work

University of Copenhagen  
Society for Danish Language  
and Literature

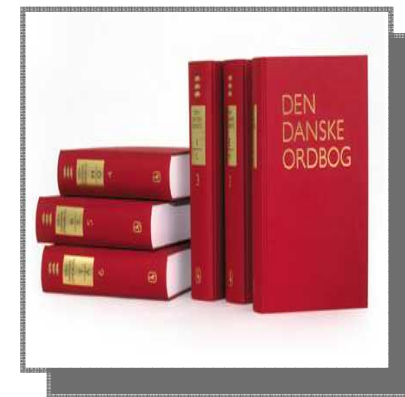


### **Danish lexical sources:**

The Danish Dictionary (DDO)  
SIMPLE-DK

### **International frameworks:**

Princeton WordNet, EuroWordNet



**Funded by:** The Danish Research Council  
for the Humanities 2004-2008

Further extensions in DK-CLARIN -> 2011



## 1st release of DanNet

### Contents:

41,000 synonym sets (synsets)  $\approx$  concepts

(Future contents: 70,000 synsets)

### Distribution on word classes:

34,000 noun synsets

6,000 verbal synsets

1,000 adjectival synsets



## 1st release of DanNet

### Coverage:

Focus on concrete nouns (27,000 synsets)

Frequency-based

### Ontological type and semantic relations:

All 41,000 synsets have an ontological type and a closests Danish hypernym

Concrete nouns have on average four semantic relations

Approx. 50 % of the material has been semi-automatically produced without further enrichment; in particular events and abstract entities

2% of the material has been validated:

Conclusion: Somewhat uneven encoding of other relations than hypernymy

(some semantic areas encoded in more depth than others)

**Still work in progress!**

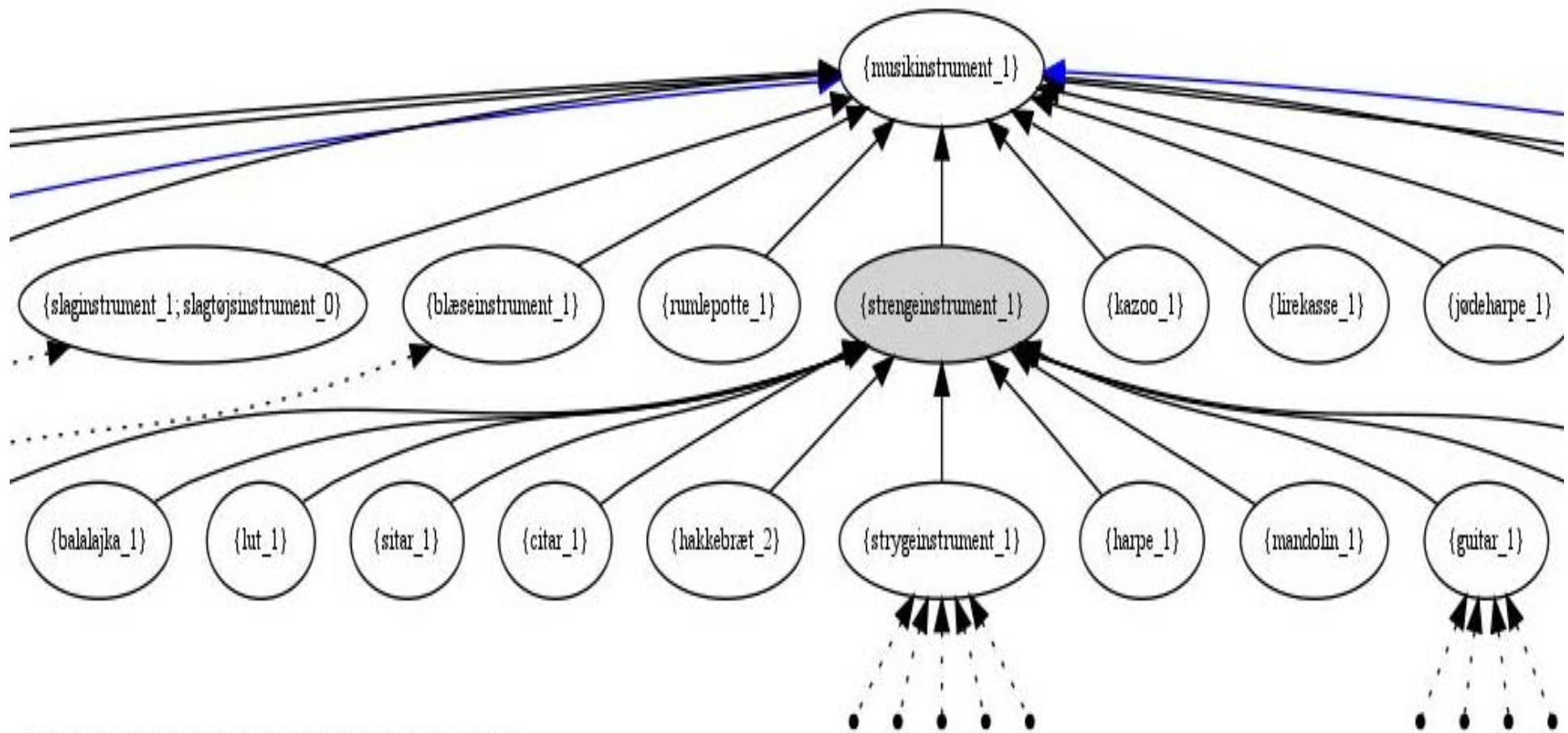


## Contents

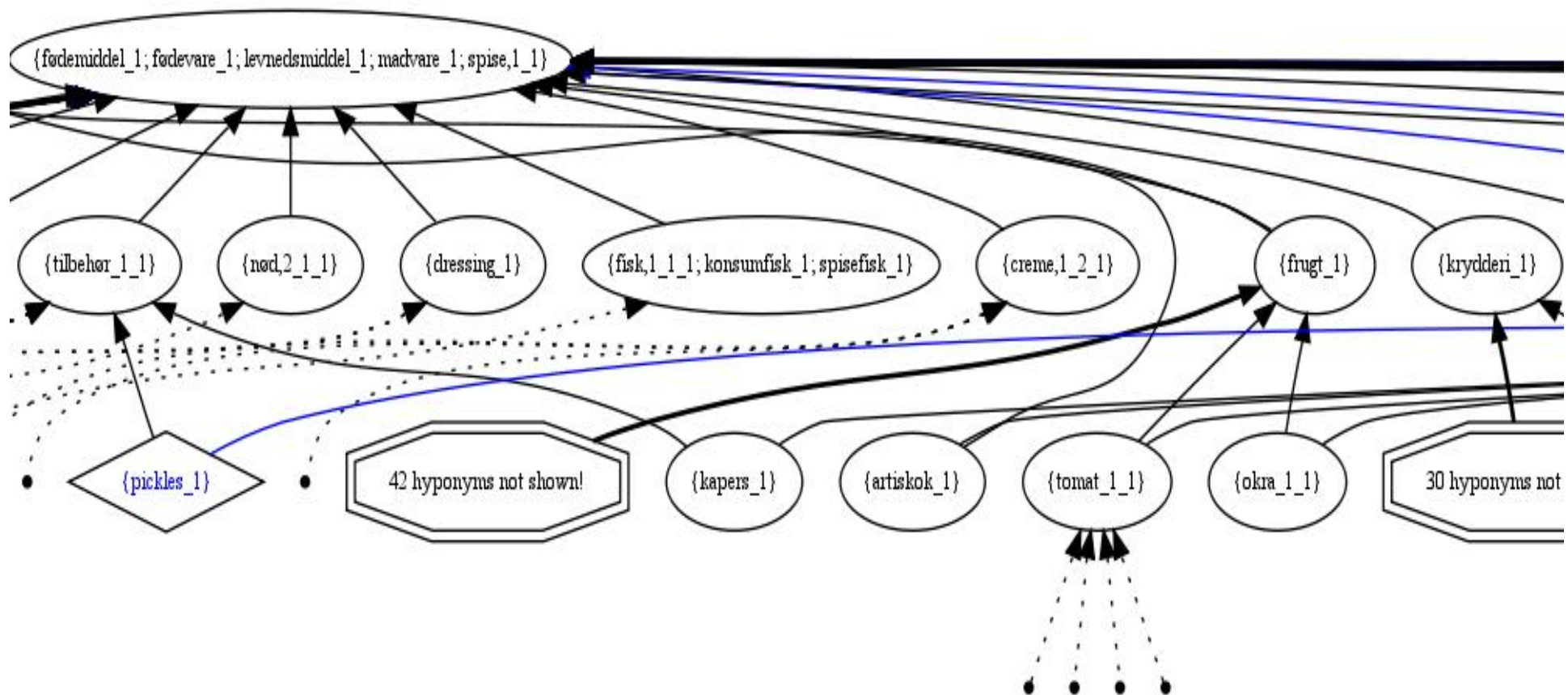
- **Taxonomical structure, simple and complex**
- **Ontological type**
- **Synonymy**
- **Polysemy**
- **Semantic features and relations**
- **Inheritance**
  
- **Getting hold of DanNet**



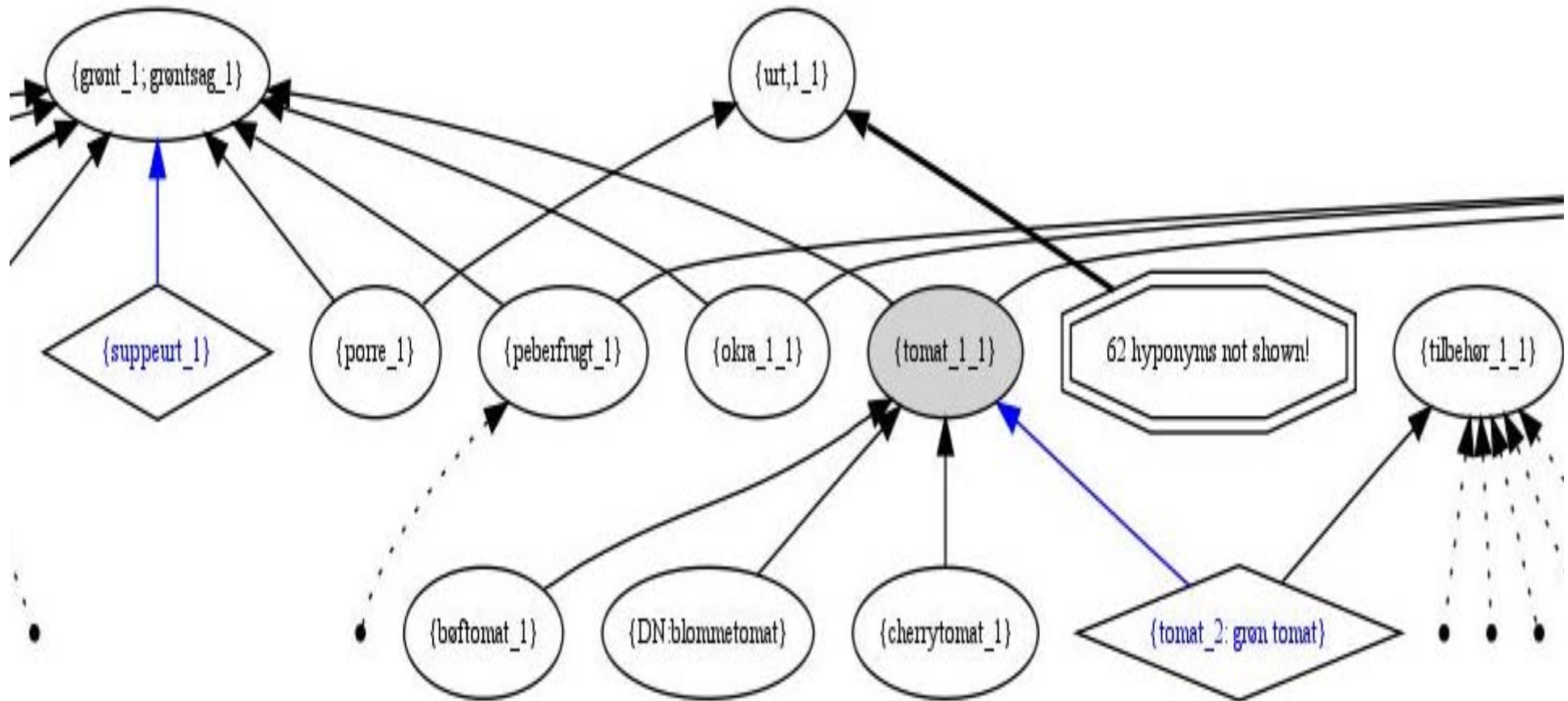
# Simple taxonomical structure: musical instruments



# Complex taxonomical structures: Foods and botanical/zooological domains



# Complex taxonomical structure: Nutrition and botanical domains



## Orthogonal vs. taxonomical: meat

- ⊖ {kød\_1\_1; kødvare\_0} (Natural+Substance): dyrs muskelvæv
  - ⊖ orthogonal hyperonymy
    - ⊕ {fars\_1} (Animal+Comestible+Part): blanding af hakket lammekød og hakket svinekød
    - ⋯ {hakkekød\_1; hakkemad\_1} (Animal+Comestible+Substance): hakket kød
    - ⋯ {hormonkød\_1} (Animal+Comestible+Substance): kød fra dyr, der er behandlet med hormoner
    - ⋯ {kødbjerg\_1} (Animal+Comestible+Object): stor mængde kød
    - ⊖ {kødstykke\_1} (Animal+Comestible+Part): b-ord
      - ⊕ taxonomic hyperonymy
    - ⊕ {pølse\_1} (Comestible+Artifact+Object): madvare i form af en pølse
    - ⋯ {saltmad\_1} (Comestible): (mad af) saltet kød og flæsk
  - ⊖ taxonomic hyperonymy
    - ⋯ Animal+Comestible+Object
    - ⊕ {fjerkræ\_1\_1} (Animal+Comestible+Object): kød fra sådanne dyr som kylling, kalkun, and
    - ⊕ {flæsk\_1; svinekød\_1} (Animal+Comestible+Substance): kød fra svine
    - ⊕ {indmad\_1} (Animal+Comestible+Part): spiselige indvolde
    - ⋯ {kalv\_1\_1; kalvekød\_1} (Animal+Comestible+Substance): kød fra kalve
    - ⊕ {lam,1\_1\_1; lammekød\_1} (Animal+Comestible+Substance): kød fra lam
    - ⋯ {okse,1\_1\_2; oksekød\_1} (Animal+Comestible+Substance): kød fra okser
    - ⋯ {snegl\_1\_1} (Animal+Comestible+Object): særlig art af snegl
    - ⊕ {vildt\_1\_1} (Animal+Comestible+Substance): kødet fra vildt



# Ontological type: EuroWordNet Ontology

## Ex: Artifact+Substance+Comestible (soup, porridge)

Top <sup>0</sup>	
1stOrderEntity <sup>1</sup>	2ndOrderEntity <sup>0</sup>
<b>Origin<sup>0</sup></b> Natural <sup>21</sup> Living <sup>30</sup> Plant <sup>18</sup> Human <sup>106</sup> Creature <sup>2</sup> Anima <sup>123</sup> Artifact <sup>144</sup>	<b>SituationType<sup>6</sup></b> Dynamic <sup>134</sup> BoundedEvent <sup>183</sup> UnboundedEvent <sup>48</sup> Static <sup>28</sup> Property <sup>61</sup> Relation <sup>38</sup>
<b>Form<sup>0</sup></b> Substance <sup>32</sup> Solid <sup>63</sup> Liquid <sup>13</sup> Gas <sup>1</sup> Object <sup>62</sup>	<b>SituationComponent<sup>0</sup></b> Cause <sup>67</sup> Agentive <sup>170</sup> Phenomenal <sup>17</sup> Stimulating <sup>25</sup> Communication <sup>50</sup> Condition <sup>62</sup> Existence <sup>27</sup> Experience <sup>43</sup> Location <sup>76</sup> Manner <sup>21</sup> Mental <sup>90</sup> Modal <sup>10</sup> Physical <sup>140</sup> Possession <sup>23</sup> Purpose <sup>137</sup> Quantity <sup>39</sup> Social <sup>102</sup> Time <sup>24</sup> Usage <sup>8</sup>
<b>Composition<sup>0</sup></b> Part <sup>86</sup> Group <sup>63</sup>	
<b>Function<sup>55</sup></b> Vehicle <sup>8</sup> Representation <sup>12</sup> MoneyRepresentation <sup>10</sup> LanguageRepresentation <sup>34</sup> ImageRepresentation <sup>9</sup> Software <sup>4</sup> Place <sup>45</sup> Occupation <sup>23</sup> Instrument <sup>18</sup> Garment <sup>3</sup> Furniture <sup>6</sup> Covering <sup>8</sup> Container <sup>12</sup> Comestible <sup>32</sup> Building <sup>13</sup>	
<b>3rdOrderEntity<sup>33</sup></b>	



Request on ontological type provides a different structure and a different extract of concepts than request on hypernym

**hypernym *Fødevarer* (food):**

*Tomat* (tomato)

*Budding* (pudding)

*Steak* (steak)

...

**Ontological type Comestible:**

Comestible+Plant+Part

*Tomat* (tomato)

Comestible+Artifact+Substance

*Budding* (pudding)

Comestible+Animal+Part

*Steak* (steak)

Comestible+Artifact+Liquids

*Lambrusco* (lambrusco)



Ontological type provides different structure and different extract of concepts than hypernym

**hypernym *Person* (person):**

all synsets under *person* such as

*Idiot* (idiot))

*Søndagsbilist* (unexperiences driver)

*Narkoselæge* (anaesthetist)

**Ontological type Human:**

Human+Object

*Idiot* (idiot))

*Søndagsbilist* (unexperiences driver)

Human+Object+Occupation

*Narkoselæge* (anaesthetist)

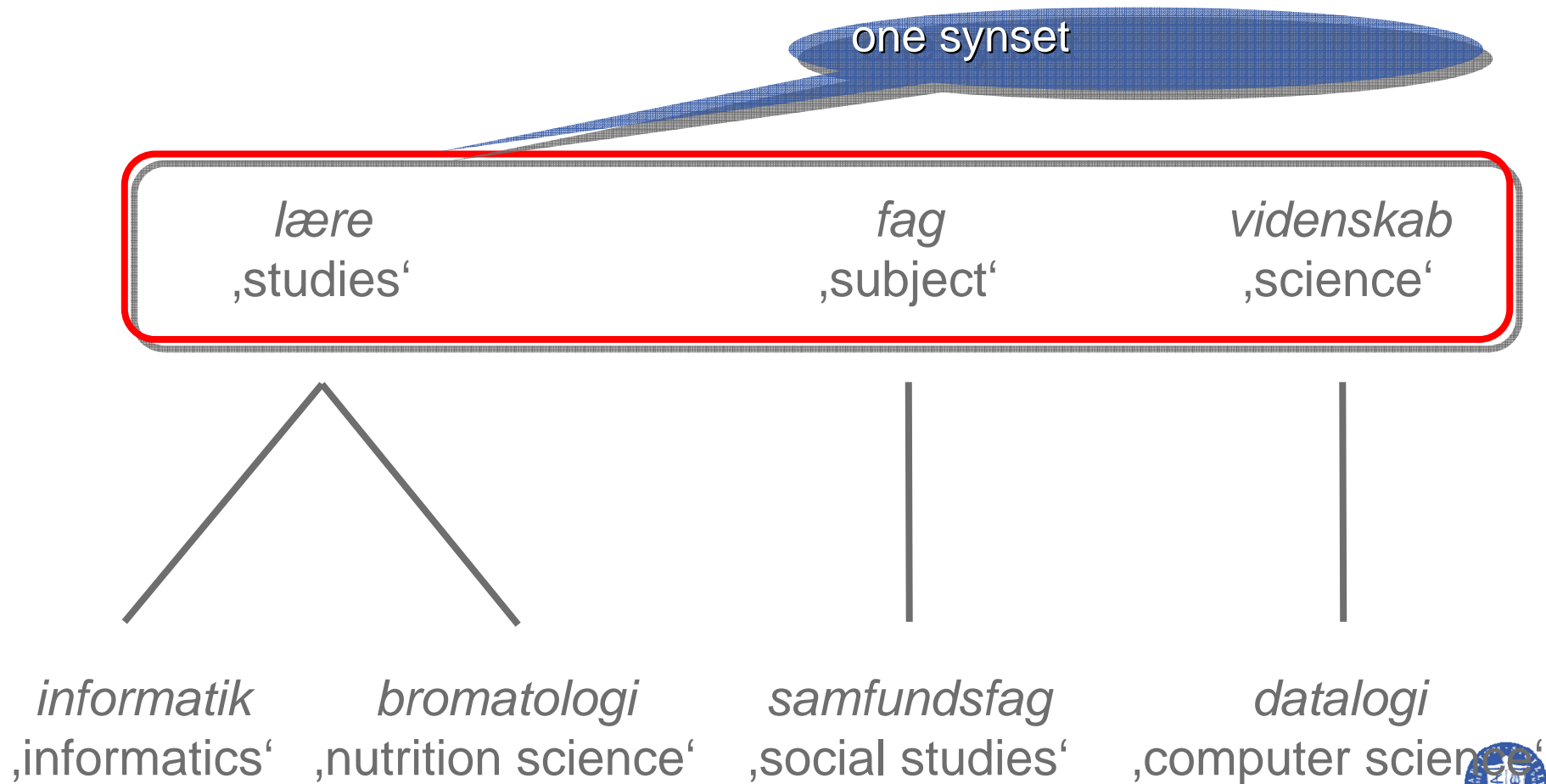
Human+Group

*Hær* (army)

*Bondestand* (peasantry)



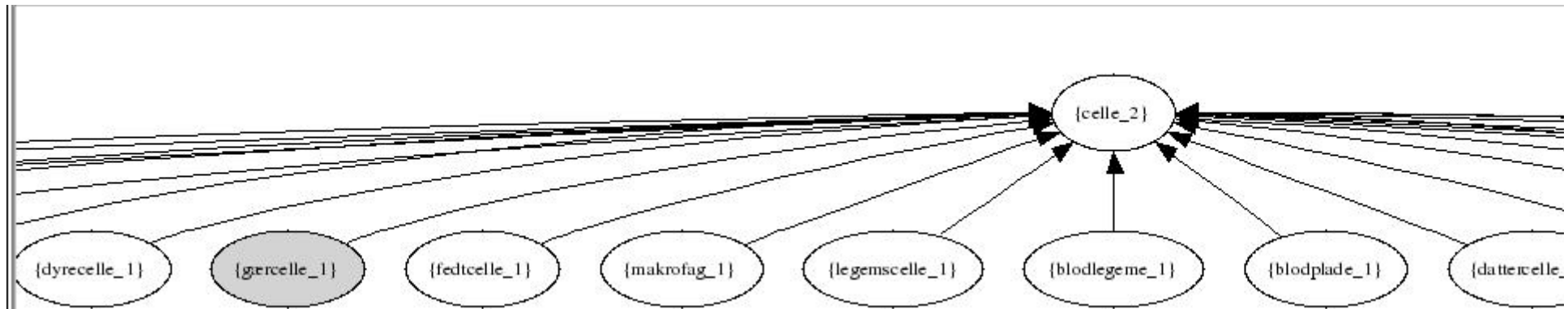
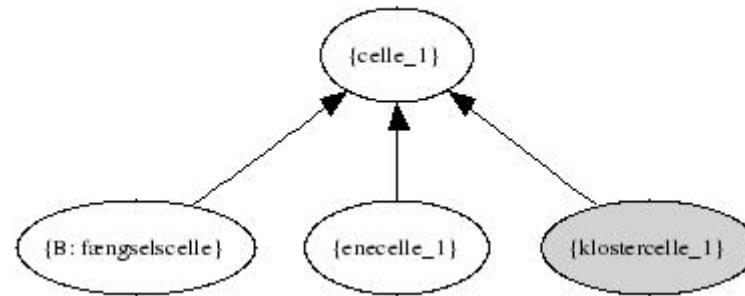
## Synonymy: several words – one concept/synset



## Polysemy: one word – different concepts

cell: room, biological cell

Complete



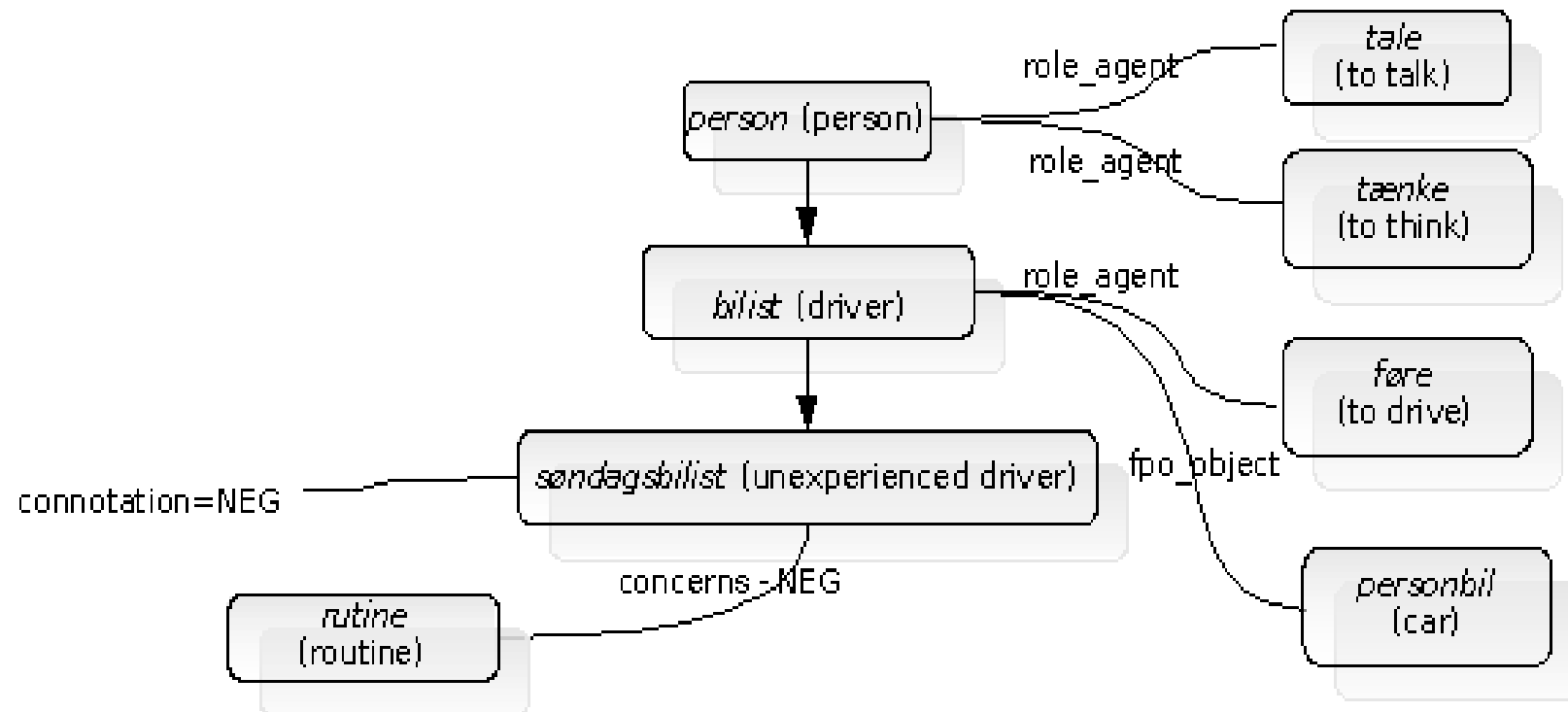
## Semantic relations in DanNet

<b>concerns</b>	<i>fodboldmål concerns sport</i>
<b>used_for</b>	<i>hammer used_for hamre</i>
<b>used_for_object</b>	<i>clipse used_for_object clips</i>
<b>made_by</b>	<i>bagværk made_by bage</i>
<b>has_holo_madeof</b> <b>has_holo_member</b> <b>has_holo_location</b> <b>has_holo_part</b>	<i>mel has_holo_madeof brød</i> <i>partimedlem has_holo_member parti</i> <i>oase has_holo_location ørken</i> <i>øje has_holo_part ansigt</i>
<b>has_hyperonym</b> <b>has_hyperonym_ortho</b>	<i>birketræ has_hyperonym træ</i> <i>vejtræ has_hyperonym_ortho træ</i>
<b>has_mero_madeof</b> <b>has_mero_member</b> <b>has_mero_part</b> <b>has_mero_location</b>	<i>brød has_mero_madeof mel</i> <i>parti has_mero_member partimedlem</i> <i>hånd has_mero_part finger</i> <i>ørken has_mero_location oase</i>
<b>role_agent</b>	<i>passager role_agent rejse</i>
<b>role_patient</b>	<i>modtager role_patient modtage</i>
<b>involved_agent</b>	<i>violin involved_agent violinist</i>
<b>involved_instrument</b>	<i>violinist involved_instrument violin</i>
<b>near_synonym</b>	<i>si near_synonym dørslag</i>
<b>xpos_near_synonym</b>	<i>behandle xpos_near_synonym behandling</i>
<b>eq_has_synonym</b>	<i>bil eq_has_synonym car</i>



# Inheritance

Ex *Søndagsbilist* (unexperienced driver): driver who lacks routine in driving and therefore drives unsafely



DanNet download: [www.wordnet.dk](http://www.wordnet.dk)

## DanNet contributors

Sanni Nimb  
Lars Trap-Jensen  
Nicolai Sørensen  
Jørg Asmussen  
Henrik Lorentzen  
Anna Braasch  
Bolette Sandford Pedersen



## Getting hold of DanNet – wordnet.dk - download

### DanNet license

(Corresponding to Princeton WordNet License)

“Permission to **use, copy, modify and distribute** this software and database and its documentation **for any purpose** and **without fee or royalty** is hereby granted provided that you agree to comply with the following copyright notice and statements...”

DanNet 1.0 Copyright 2009 by University of Copenhagen and the Danish Society of Language and Literature. All rights reserved



## Available format ([www.wordnet.dk](http://www.wordnet.dk))

W3C-standard for wordnets:

Xml

Owl/rdf instance model

Demo can be viewed in Protegé 4.0



The screenshot shows the Protege OWL editor interface for the file 'DanNet\_full.owl'. The main window is divided into several panes:

- Active Ontology:** Shows the current ontology being edited.
- Entities:** A list of individuals, with '{rustvogn\_1}' selected. Other individuals include {højtalerovogn\_1}, {kølevogn\_1}, {ladvogn\_1}, {landrover\_1}, {ministerbil\_1}, {pejlevogn\_1}, {postvogn\_2}, {reklamevogn\_1}, {reportagevogn\_1}, {riskoger\_2}, {rumlekaske\_1}, {sardindåse\_1\_1}, {servicevogn\_1}, {servicevogn\_2}, {skolevogn\_1}, {skramlekaske\_1}, {slagskib\_1\_1}, {standardvogn\_1}, {standardvogn\_2}, {startvogn\_1}, {stockcar\_1}, {terrænbil\_1}, {testvogn\_1}, {udlejningsbil\_1}, {veteranbil\_1}, {økonomibil\_1}, {øse\_1\_2}, {DN:lånebil}, {trolleybus\_1}, {blokvogn\_1\_1}, {forvogn\_2}, {miljøbil\_1}, {trækker\_1}, and {hybridbus\_1}.
- Individual Annotations:** Shows annotations for the selected individual, including a 'label' with the value '{rustvogn\_1}'.
- Description:** Shows the type 'NounSynset' for the individual.
- Property assertions:** Lists various properties and their values for the individual, such as 'usedFor' (transportere\_1), 'hyponymOf' (automobil\_1, bil\_1, dyt\_2\_1, karet\_1\_2, slæde\_1\_3, vogn\_2), 'locationMeronymOf' (land\_3, landjord\_1), and 'synsetId' (36009).

