



# Visualizing sloWNet

Darja Fišer

University of Ljubljana

Jernej Novak

University of Maribor

- What is sloWNet
- Motivation & Background for this research
- Presentation of sloWTool
- Conclusions
- Future plans

# Presentation outline

---

- Based on Princeton WordNet (Fellbaum 1998)
  - words grouped into sets of synonyms, each expressing a concept
  - synsets are interlinked with semantic and lexical relations
  - 118,000 synsets, 157,000 literals, 207,000 lit-syn pairs
- Developed automatically from heterogeneous resources
  - bilingual dictionary
  - parallel corpus
  - Wikipedia
  - partially manually revised
  - used for semantic annotation of jos100k corpus
- sloWNet 3.0:
  - 43,000 synsets, 39,700 literals, 82,700 syn-lit pairs
  - average synset length 1.92, average polysemy 2.07

# sloWNet: WordNet for Slovene

---

## dictionaries

- word-based
- semantic info implicit
- for human users

dictionary-writing  
systems  
(e.g. TshwaneLex )

## semantic lexicons

- concept-based
- explicit encoding of semantic info
- still language-oriented
- intended for human as well as machine use

wordnet browsers,  
editors, visualizers  
(which tool to use?)

## ontologies

- highly formalized
- compiled for reasoning in AI tasks

ontology-management  
tools  
(e.g. Protégé)

# Motivation & Background

- available tools:
  - Princeton WordNet: Princeton WordNet Browser
  - EuroWordNet: Polaris & Periscope
  - BalkaNet: DEBVisDic
- ideal tool:
  - freely available, platform-independent, on-line
  - all-in-one tool (browsing, editing & visualization)
  - support for standardized formats (e.g. LMF)
  - support for multilingual scenarios
  - easy integration of third-party resources (e.g. domains, coarse-grained sense clusters, images)

# Analysis of existing wn tools

---

# sloWTool

select language

Slovenian ▾

Link

main menu



Search here, empty for random ...

Number of hits: 0

no. of results

search field

copy URL

**prst**

Slovenian ▾

[Link](#)

Number of hits: 9

POS: Noun ID: eng-30-05566097-n BCS: 2 DOMAIN: anatomy CLUSTERID: finger3

[meta info](#)SYNONYM (SLV): **daktil, prst**[synonyms](#)SYNONYM (ENG): **dactyl, digit**[definition](#)DEFINITION: *a finger or toe in human beings or corresponding body part in other vertebrates*→ [ENG\_DERIVATIVE]: **digitalen, digital**→ [HOLO\_PART]: **vretenčar, craniate, vertebrate**→ [HYPERNYM]: **ud, del, član, ud, okončina, končina, ekstremiteta, member, appendage, extremity**

POS: Noun ID: eng-30-05559908-n BCS: 2 DOMAIN: anatomy CLUSTERID: life10

SYNONYM (SLV): **član, del, ekstremiteta, končina, okončina, ud, ud**[example](#)SYNONYM (ENG): **appendage, extremity, member**DEFINITION: *an external body part that projects from the body*USAGE: *it is important to keep the extremities warm*→ [HYPERNYM]: **zunanjí del telesa, external body part**

STAMP: darja 2008-01-01 00:00:00

[edit info](#)

# Simple search

Fields

Domains

Language: Slovenian

Literals: prs\*

wildcards \* ?

Part of speech:  Adjective  Verb  Noun  Adverb

Synset ID:

Definitions:

Examples:

Stamps:

Empty

BCS:  1  2  3  Empty

Search

Clear

# Advanced search



Login

Register

Name:

Surname:

UserName:

Email:

Password:

Retype password:

register & login

POS: Noun ID: eng-30-05297523-n BCS: 1 DOMAIN: anatomy

SYNONYM (SLV): **organ, orgle**



correct errors

SYNONYM (ENG): **organ**

DEFINITION: *a fully differentiated structural and functional unit in an animal that is specialized for some particular function*

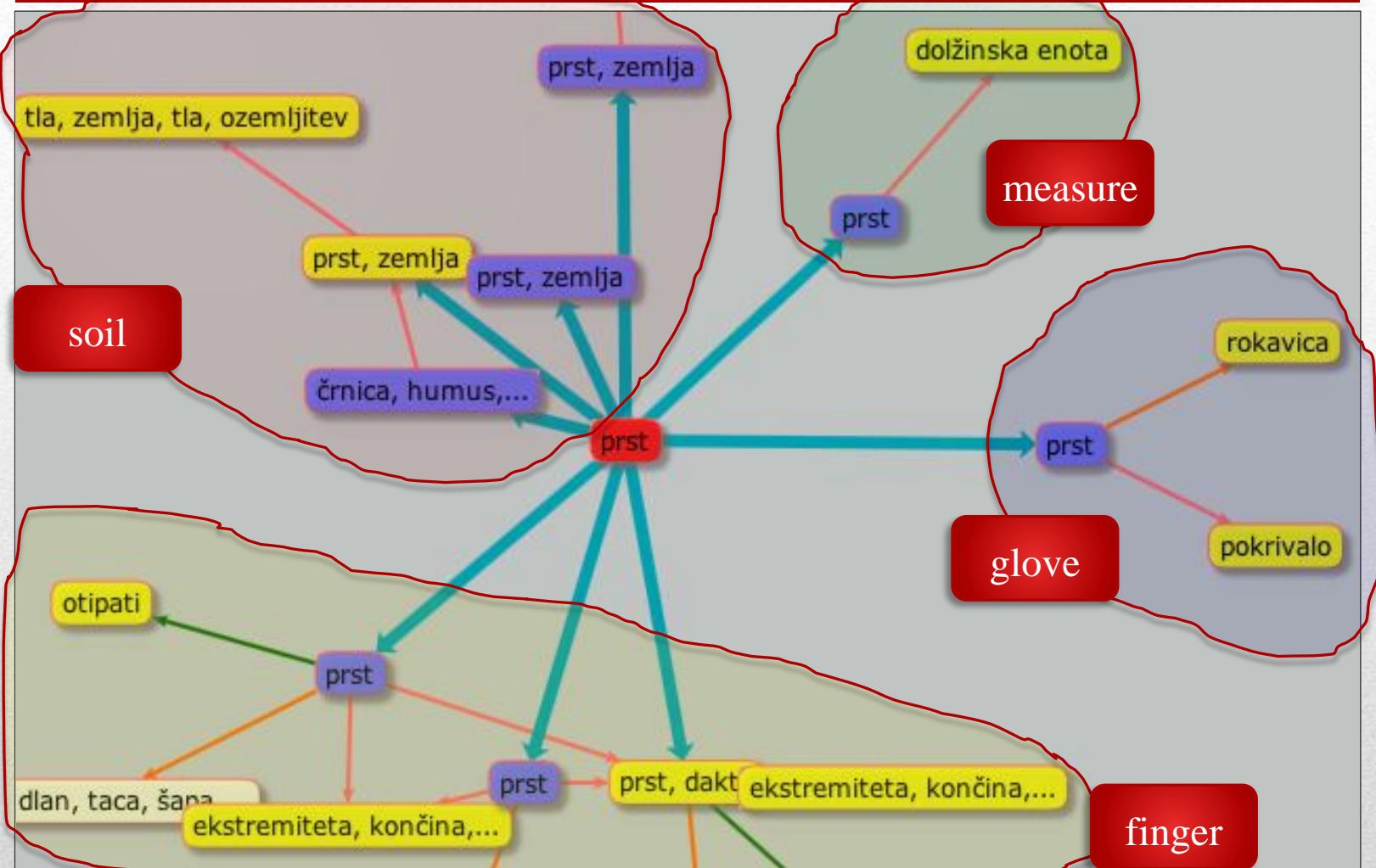
→ [ENG\_DERIVATIVE]: **organski, organic**

→ [HYPERNYM]: **del telesa, body part**

STAMP: darja 2008-01-01 00:00:00

# Editing





# Visualization



- Advantages of sloWTool:
  - all-in-one browser, editor & visualisaton tool for wordnets
  - can be used for wordnets in other languages as well as for other similar databases
  - allows simple or advanced search & view settings
- External resources integrated in sloWTool:
  - semantically annotated corpus josSENSE (Fišer&Erjavec 2010)
  - WordNet Domains Hierarchy (Bentivogli et al. 2004)
  - Coarse-grained Sense Clusters (Navigli 2006)
- availability of sloWTool:
  - for browsing: <http://kista.asuscomm.com/search>
  - for download: <https://launchpad.net/slowtool>

# Conclusions

---

- sloWTool enhancements:
  - enhance the graph-based visualization module
  - extend sloWTool to allow semantic annotation of corpora
  - extend sloWTool for crowd-sourcing sloWNet refinement
- integration of external resources:
  - add wordnets for other languages (e.g. WOLF)
  - add Princeton semantically annotated glosses
  - integrate with ImageNet database
  - add links to GeoWordNet

# Future plans

---

# Visualizing slowNet



**Thank you for your attention!**

general questions:

[darja.fiser@ff.uni-lj.si](mailto:darja.fiser@ff.uni-lj.si)

technical questions:

[jernej.novak1@uni-mb.si](mailto:jernej.novak1@uni-mb.si)

---