INFORUM 2016: 22nd Annual Conference on Professional Information Resources, Prague 24–25 May 2016



University of Belgrade - University Library "Svetozar Markovic"

Multilingual Vocabularies in Open Access:

Semantic Network WordNet

MSc Sanja Antonic,

MSc Oja Krinulovic

MSc Mile Stijepovic

University of Belgrade, University library "Svetozar Marković"



WordNet is lexico-semantic network created on Princeton University in 1985 and constantly developing, well-known like Princeton WordNet (PWN).

PRINCETON WORDNET CONTAINS INFORMATION ABOUT NOUNS, VERBS,

ADJECTIVES AND ADVERBS IN ENGLISH. THEY ARE GROUPED INTO SETS OF

SYNONYMS – SYNSETS, AND EVERY SYNSET REPRESENTS A PARTICULAR CONCEPT.

THE MAIN SEMANTIC RELATION AMONG SYNSETS IS SYNONYMY AND THERE ARE

ALSO HYPONYMY, ANTONYMY, MERONYMY ETC.

THE NEXT PHASE WAS MULTULINGUALITY PROJECT EUROWORDNET (EWN)





WordNet Search - 3.1 - WordNet home page - Glossary - Help Word to search for: cat Search WordNet Display Options: Show all Change (Select option to change) Hide Example Sentences Hide Glosses **Show Frequency Counts** Show Database Locations Show Lexical File Info Show Lexical File Numbers Show Sense Keys Show Sense Numbers Show all Hide all





WordNet Search - 3.1

- WordNet home page - Glossary - Help

Word to search for: cat Search WordNet

Display Options: (Select option to change) ▼ Change

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

Display options for sense: (gloss) "an example sentence"

Noun

- S: (n) cat, true cat (feline mammal usually having thick soft fur and no ability to roar: domestic cats; wildcats)
- S: (n) guy, cat, hombre, bozo, sod (an informal term for a youth or man) "a nice guy"; "the guy's only doing it for some doll"; "the poor sod couldn't even buy a drink"
- S: (n) cat (a spiteful woman gossip) "what a cat she is!"
- S: (n) kat, khat, qat, quat, cat, Arabian tea, African tea (the leaves of the shrub Catha edulis which are chewed like tobacco or used to make tea; has the effect of a euphoric stimulant) "in Yemen kat is used daily by 85% of adults"
- S: (n) cat-o'-nine-tails, cat (a whip with nine knotted cords) "British sailors feared the cat"
- S: (n) <u>Caterpillar</u>, cat (a large tracked vehicle that is propelled by two endless metal belts; frequently used for moving earth in construction and farm work)
- S: (n) big cat, cat (any of several large cats typically able to roar and living in the wild)
- S: (n) computerized tomography, computed tomography, CT, computerized axial tomography, computed axial tomography. CAT (a method of examining body)







Home

Consortium

Resources
Publications
Links

Steering committee
Workplan and packages

Wordnets

WordNet

This is the status of the final version of the monolingual Wordnets produced by the end of the project:

BalkaNet - Design and Development of a Multilingual Balkan

B

	Bulgarian	Czech	Greek	Romanian	Turkish	Serbian
Synsets	21441	28456	18461	19839	14626	8059
Nouns	14174	21009	14426	13345	11059	5919
Verbs	4169	5155	3402	4808	2725	1803
Adjectives	3088	2128	617	852	802	324
Adverbs	9	164	16	834	40	13
Literals	44956	43918	24366	33690	20310	13295
Literal/Synset	2.1	1.54	1.33	1.7	1.39	1.65
BC1s	1218	1218	1218	1218	1220	1219
BC2s	3471	3471	3462	3471	3479	3469
BC3s	3827	3827	3825	3827	3794	1369
Domain Specific Synsets	2065	304	238	286	300	305
Balkan Specific Synsets	220	257	309	151	103	117
Language Specific Synsets	116	257	52	545	204	206
Language Internal Relations	28599	25683	24368	25885	19834	12787

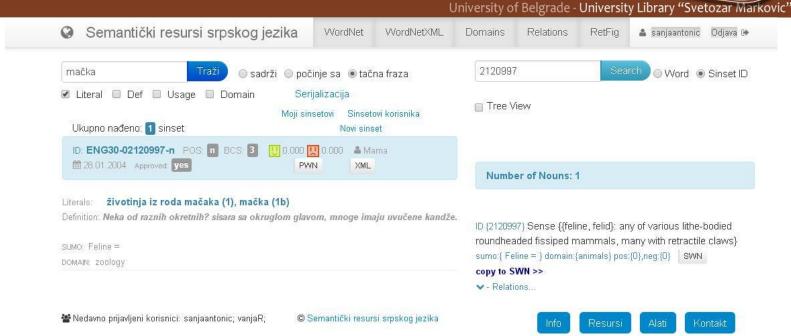




University of Belgrade - University Library "Svetozar Markovic" Semantički resursi srpskog jezika WordNet WordNetXML Domains Relations RetFig 📤 sanjaantonic Odjava 🕩 832092 Pretražite srpski WordNet. ● sadrži ○ počinje sa ○ tačna fraza WordSinset ID ☑ Literal ☐ Def ☐ Usage ☐ Domain Serijalizacija Tree View Moji sinsetovi Sinsetovi korisnika Ukupno nađeno: 429 sinsetova Novi sinset ID: ENG30-00832092-n POS: n BCS: 🕎 0.000 🕎 0.000 - 🚨 sanjaantonic ## 08.08.2015 Approved: yes PWN XML Izmeni Number of Nouns: 1 sinset kardiografija (1), elektrokardiografija (1), EKG (1) ID (832092) Sense {{cardiography, electrocardiography}: Definition: Dijagnostička procedura koja prati rad srca. Eletrokardiografijom se korišćenjem diagnostic procedure consisting of recording the activity of the aparata elektrokardiograma prate električni impulsi srca i dobija crtež elektrokardiograf - EKG. heart electronically with a cardiograph (and producing a cardiogram)} SUMO: TherapeuticProcess + sumo:{ TherapeuticProcess + } domain:{medicine} pos:{0},neg:{0} DOMAIN: medicine copy to SWN >> ID: ENG30-00613685-n POS: 10 BCS: U 0.000 Q 0.000 a sanjaantonic Relations... ## 08.08.2015 Approved: yes PWN XML Izmeni sinset biljna medicina (1) Definition: Upotreba lekovitih biljaka u sprečavanju, lečenju bolesti ili promovisanju zdravlja SUMO: Position + DOMAIN: pharmacy ID: ENG30-00086077-v POS: V BCS: U 0.000 Q 0.000 a sanjaantonic ## 29.06.2015 Approved: yes XML Izmeni







Japanese WordNet



University of Belgrade - University Library "Svetozar Markovic"

Japanese WordNet

Introduction

News

Release & Downloads

Illustrations

References

Related Projects

Search WordNet

→ <u>日本語 (Japanese)</u>

Francis Bond

bond@ieee.org>

NTU Computational Linguistics Lab

Linguistics and Multilingual Studies
Nanyang Technological University

Results for «ネコ» (jpn)

02121620-n (18) にゃんにゃん,猫,ネコ, キャット

cat, true cat

通常、厚く柔らかい毛皮を持ち、吠えることのできないネコ

乳類:家ネコ;ヤマネコ

Search WN ネコ Langs: 日本語▼ 英語 ▼

More detail about the Japanese Wordnet (1.1)

This project is now integrated in the Extended Open Multilingual Wordnet (1.1)

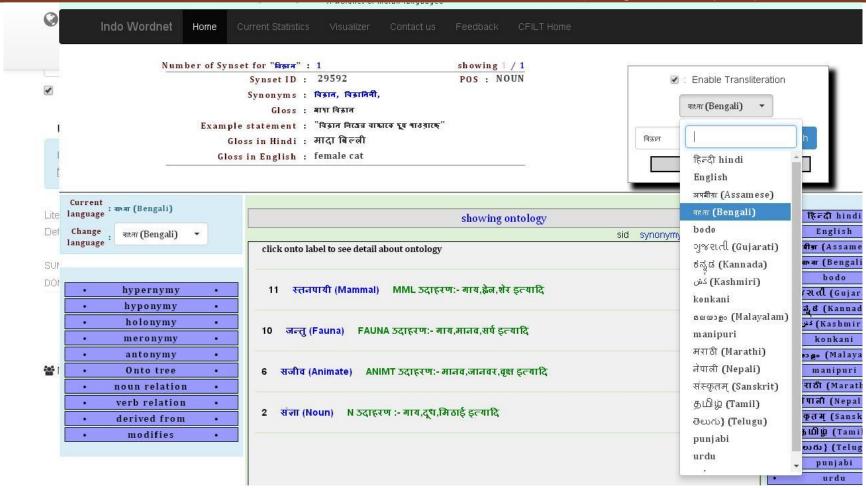
Maintainer: Francis Bond Sond@ieee.org

(0.001

IndoWordNet



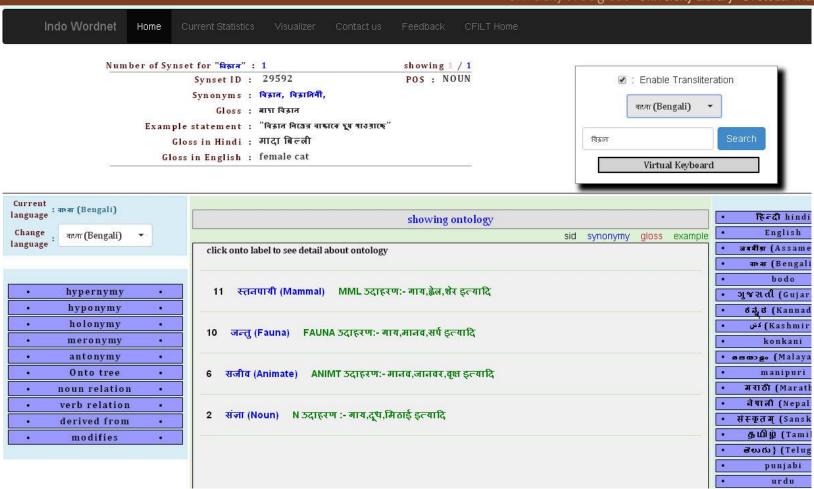
University of Belgrade - University Library "Svetozar Markovic"



IndoWordNet



University of Belgrade - University Library "Svetozar Markovic"





Thank You

antonic@unilib.rs okrinulovic@unilib.rs stijepovic@unilib.rs

www.unilib.rs