

# Exploring Onto.PT



U C FCTUC

Hugo Gonalo Oliveira<sup>a</sup>, Leticia Ant3n P3rez & Paulo Gomes  
 hroliv@dei.uc.pt, leticiaap86@gmail.com,  
 pgomes@dei.uc.pt

Cognitive & Media Systems Group  
 CISUC, University of Coimbra, Portugal



<sup>a</sup>Hugo Gonalo Oliveira is supported by FCT, grant SFRH/BD/44955/2008, co-funded by FSE.

## Introduction

- Public **lexical ontology** for Portuguese
  - Freely available from <http://ontopt.dei.uc.pt/>
- Modelled after Princeton WordNet
- Populated **automatically** after exploiting public lexical resources
  - Semantic relations, connecting lemmas, extracted from dictionaries [1]
  - Attachment of synonymy relations to suitable synsets in a thesaurus and discovery of new synsets in the unattached synonymy relations [2]
  - Assignment of lemmas in relations to the most suitable synsets [3]

## Integrated resources

- Relations from **dictionaries** – CART3O [1]
  - PAPEL [4] (v.3.0)
  - Dicion3rio Aberto [5]
  - Wiktionary.PT<sup>a</sup>
- Handcrafted **thesauri**
  - TeP 2.0 [6]
  - OpenThesaurus.PT<sup>b</sup>

## Current contents (v.0.3.1)

- ≈160,000 **lexical items**
- ≈110,000 **synsets**
  - Noun (64,865)
  - Verb (25,342)
  - Adjective (17,952)
  - Adverb (1,993)
- ≈170,000 instances of **semantic relations**
  - relation types from PAPEL 3.0

Group	Relation	Instances
Hypernym	<i>n</i> hiperonimoDe <i>n</i>	83,552
Part	<i>n</i> parteDe <i>n</i>	3,672
	<i>n</i> parteDeAlgoComProp <i>adj</i> <i>adj</i> propDeAlgoParteDe <i>n</i>	91
Member	<i>n</i> membroDe <i>n</i>	5,847
	<i>n</i> membroDeAlgoComProp <i>adj</i> <i>adj</i> propDeAlgoMembroDe <i>n</i>	106
Contains	<i>n</i> contidoEm <i>n</i>	355
	<i>n</i> contidoEmAlgoComProp <i>adj</i>	264
Material	<i>n</i> materialDe <i>n</i>	835
Causation	<i>n</i> causadorDe <i>n</i>	1,347
	<i>n</i> causadorDeAlgoComProp <i>adj</i>	26
	<i>adj</i> propDeAlgoQueCausa <i>n</i>	619
	<i>n</i> causadorDaAccao <i>v</i> <i>v</i> accaoQueCausa <i>n</i>	56
Place	<i>n</i> localOrigemDe <i>n</i>	8,052
	<i>n</i> localOrigemDe <i>n</i>	1,293
Antonym	<i>adj</i> antonimoAdjDe <i>adj</i>	538
Producer	<i>n</i> produtorDe <i>n</i>	1,718
	<i>n</i> produtorDeAlgoComProp <i>adj</i> <i>adj</i> propDeAlgoProdutorDe <i>n</i>	88
Purpose	<i>n</i> fazSeCom <i>n</i>	529
	<i>n</i> fazSeComAlgoComProp <i>adj</i>	6,551
	<i>v</i> finalidadeDe <i>n</i> <i>v</i> finalidadeDeAlgoComProp <i>adj</i>	79
Quality	<i>n</i> temQualidade <i>n</i>	7,271
	<i>n</i> devidoAQualidade <i>adj</i>	322
State	<i>n</i> temEstado <i>n</i>	934
	<i>n</i> devidoAEstado <i>adj</i>	1,059
Manner	<i>adv</i> maneiraPorMeioDe <i>n</i>	327
	<i>adv</i> maneiraComProp <i>adj</i>	197
Manner without	<i>adv</i> maneiraSem <i>n</i>	1,833
	<i>adv</i> maneiraSemAccao <i>v</i>	1,561
Property	<i>adj</i> dizSeSobre <i>n</i>	216
	<i>adj</i> dizSeDoQue <i>v</i>	14
		9,145
		25,014

## OntoBusca interface

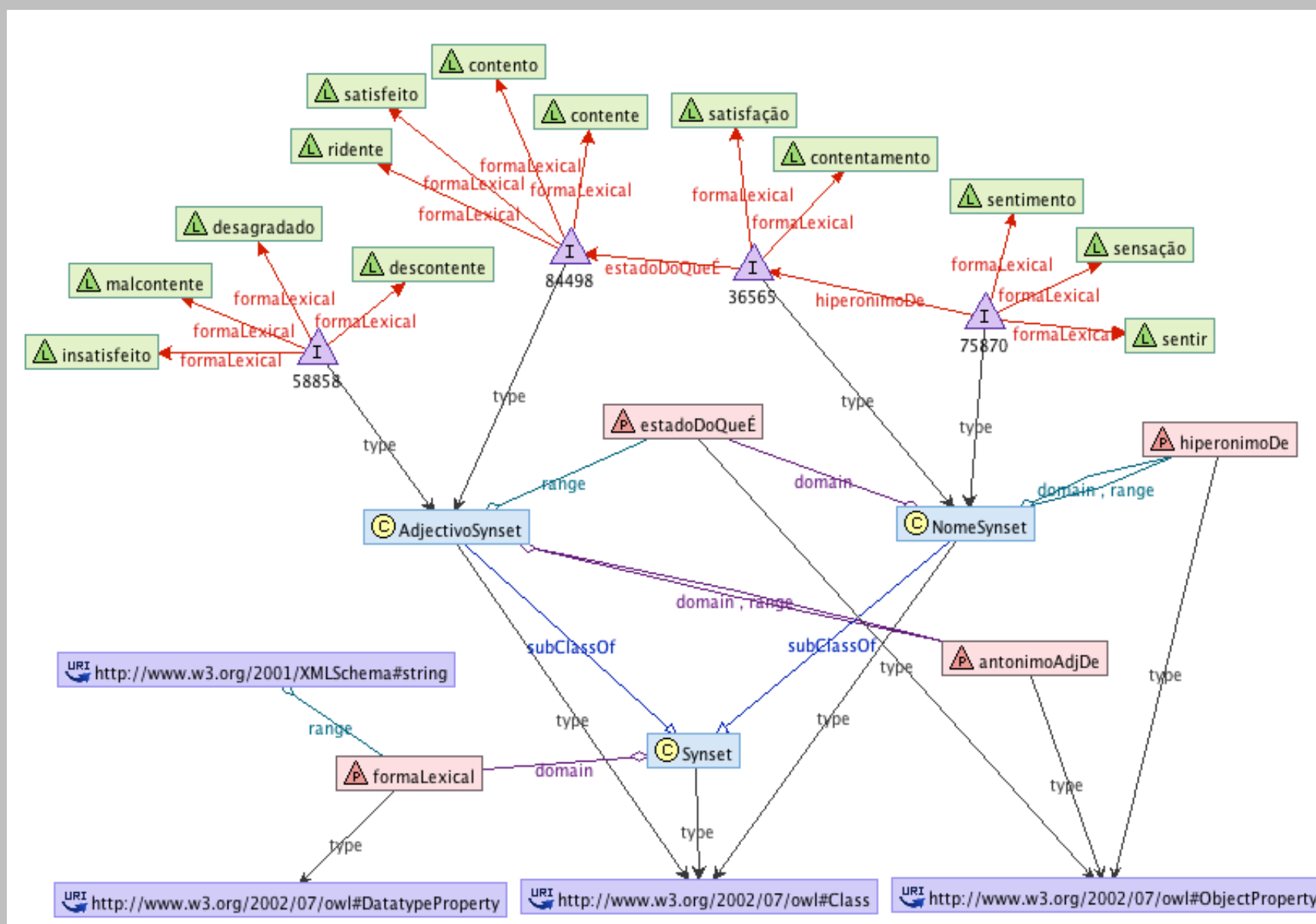
- Retrieves synsets containing a queried lexical item
- Expands relations where they are involved
- Cloud with most frequent searches



- S:** (n) **hospital**, nosoc3mio, nosoc3mio, espirital
  - hiperonimoDe
    - S:** (n) manic3mio, riilhafoles
    - S:** (n) lazareto
    - S:** (n) leprosarria, gafaria
    - S:** (n) leproc3mio
  - hiponimoDe
    - S:** (n) empresa, casa, estabelecimento, firma
    - S:** (n) obra, casa, edif3cio, pr3dio, bloco, edificac3o, herdade, edificamento, im3vel
    - S:** (n) edif3cio, pr3dio, aranha-c3u
  - referidoPeloQueE
    - S:** (adj) hospital3rio, hospitalar, nosoc3mico, nosocomial, nosoc3mico
  - temParte
    - S:** (n) enferm3ria
      - parteDe
      - hiperonimoDe
        - S:** (n) ambulat3rio
      - servePara
        - S:** (n) paciente, enfermo, doente
      - hiponimoDe
- S:** (adj) **hospital**, benfeitor, virtuoso, humano, santo, bom, favor3vel, bendito, piedoso, misericordioso, propenso, ben3volo, patriarcal, indulgente, clemente, benigno, ben3fico, benevolente, generoso, bondoso, obsequioso, caridoso, almo, caritativo, bondadoso, benfazente, agatoide, bondoso
- S:** (adj) **hospital**, esmoler, esmolador, caridoso, caritativo

## RDF/OWL model

- Based on the W3C WordNet RDF/OWL Basic [7].



## References

- H. Gonalo Oliveira, L. Ant3n P3rez, H. Costa, and P. Gomes, "Uma rede l3xico-sem3ntica de grandes dimens3es para o portugu3s, extraída a partir de dicion3rios electr3nicos," *Linguam3tica*, vol. 3, pp. 23–38, December 2011.
- H. Gonalo Oliveira and P. Gomes, "Automatically enriching a thesaurus with information from dictionaries," in *Proc 15th Portuguese Conference on Artificial Intelligence (EPIA 2011)*, vol. 7026 of LNCS, pp. 462–475, Springer, October 2011.
- H. Gonalo Oliveira and P. Gomes, "Ontologising relational triples into a portuguese thesaurus," in *Proc. 15th Portuguese Conference on Artificial Intelligence, EPIA 2011*, (Lisbon, Portugal), pp. 803–817, APPIA, October 2011.
- H. Gonalo Oliveira, D. Santos, and P. Gomes, "Extrac3o de relat3es sem3nticas entre palavras a partir de um dicion3rio: o PAPEL e sua avalia3o," *Linguam3tica*, vol. 2, pp. 77–93, May 2010.
- A. Sim3es, A. I. Sanrom3n, and J. J. ao Almeida, "Dicion3rio-aberto: A source of resources for the portuguese language processing," in *Proc. of PROPOR 2012*, vol. 7243 of LNCS, (Coimbra, Portugal), pp. 121–127, Springer, April 2012.
- E. G. Maziero, T. A. Pardo, A. Di Felippo, and B. C. Dias-da-Silva, "A Base de Dados Lexical e a Interface Web do TeP 2.0 - Thesaurus Eletr3nico para o Portugu3s do Brasil," in *VI TIL*, pp. 390–392, 2008.
- M. van Assem, A. Gangemi, and G. Schreiber, "RDF/OWL representation of WordNet," working draft, W3C, June 2006.

<sup>a</sup><http://pt.wiktionary.org/>

<sup>b</sup><http://openthesaurus.caixamagica.pt/>