

# Hontology

## 1. BASIC INFORMATION

1. *Lexicon type (wordform, explanatory, terminological lexicon, wordnet, etc.)*

Hontology (H stands for hotel, hostel and hostel) (available at <http://ontolp.inf.pucrs.br/Recursos/downloads-Hontology.php>) is a new multilingual ontology for the accommodation sector freely available, containing 282 concepts categorized into 16 top-level concepts.

The concepts of other vocabularies such as Dbpedia.org and Schema.org. were reused and new concepts and properties from online reviews (e.g. booking.com or realtravel.com) with examples of multilingual texts provided by products and services consumers were integrated and manually annotated with different labels, through terms correlation, rules (lexical patterns) and synonyms, by experts, in four languages: English, Portuguese, Spanish and French. For more information about the Hontology development, see Chaves and Trojahn (2010).

Hontology is useful for a wide range of applications within the accommodation sector, including ontology-based information extraction, opinion mining (sentiment analysis) (see Pang and Lee, 2008), text annotation and information visualization.

Hontology contents:

Metrics	Values
Number of Concepts	282
Number of Object Properties	8
Number of Data Properties	31
<b>Concept Axioms</b>	
Subconcept axioms	272
Equivalent concepts axioms	4
Disjoint concepts axioms	93
<b>Object Property Axioms</b>	
Functional object property axioms	5
Object property domain axioms	9
Object property range axioms	9
<b>Data Property Axioms</b>	

Functional data property axioms	12
Object data domain axioms	17
Object data range axioms	1

### *1.2 Representation of the lexicon (flat files, database, markup)*

Applications using Hontology explore different levels of formality. Some of them might need to perform reasoning, while other work just with a flat list of concepts and properties. In order to satisfy these needs, the ontology is available in OWL, RDF and XML formats.

### *1.3 Character encoding*

The characters are in UTF-8 code.

## **2. ADMINISTRATIVE INFORMATION**

### *2.1 Contact person (name, address, affiliation, position, telephone, fax, e-mail)*

Name: Marcirio Chaves  
Address: Universidade Atlântica  
Fábrica da Pólvora de Barcarena  
2730-036 BARCARENA  
Portugal  
Position: Assistant professor  
Affiliation: Universidade Atlântica  
Telephone: +351 21 439 82 00 (Ext.: 217)  
Fax: +351 21 430 25 73  
E-mail: marcirioc@uatlantica.pt

### *2.2 Delivery medium (if relevant; description of the content of each piece of medium)*

The resource will be available on the META-SHARE platform.

### *1. Copyright statement and information on IPR*

Hontology is a free license-based for research purposes only, with attribution and distribution allowed under the original license: CC-BY-NC-SA 2.5 (license at <http://creativecommons.org/licenses/by-nc-sa/2.5/br/>).

## **3. TECHNICAL INFORMATION**

### *1. Directories and files*

The archive contains one OWL file with the complete structure/grammar of the ontology.

### 3.2 Data structure of an entry

Considering that this ontology is available in OWL, RDF and XML formats, following a class sample for Boat Hotel entry in OWL format which contains more information, and a sample of terms correlation between two concepts:

```
<AnnotationAssertion>
  <AnnotationProperty abbreviatedIRI="rdfs:comment"/>
  <AbbreviatedIRI>accommodation-04042012:Botel</AbbreviatedIRI>
  <Literal xml:lang="en" datatypeIRI="&rdf;PlainLiteral">Boat
Hotel</Literal>
</AnnotationAssertion>
```

---

```
<DisjointClasses>
  <Class abbreviatedIRI="hotel-15072010:Apartment"/>
  <Class abbreviatedIRI="hotel-15072010:GuestHouse"/>
</DisjointClasses>
```

### 3.3 Lexicon size (nmb. of lexical items, KB occupied on disk)

The Hontology is composed by 282 concepts with 16.5 KB compressed (229.1 KB uncompressed) for disk storage.

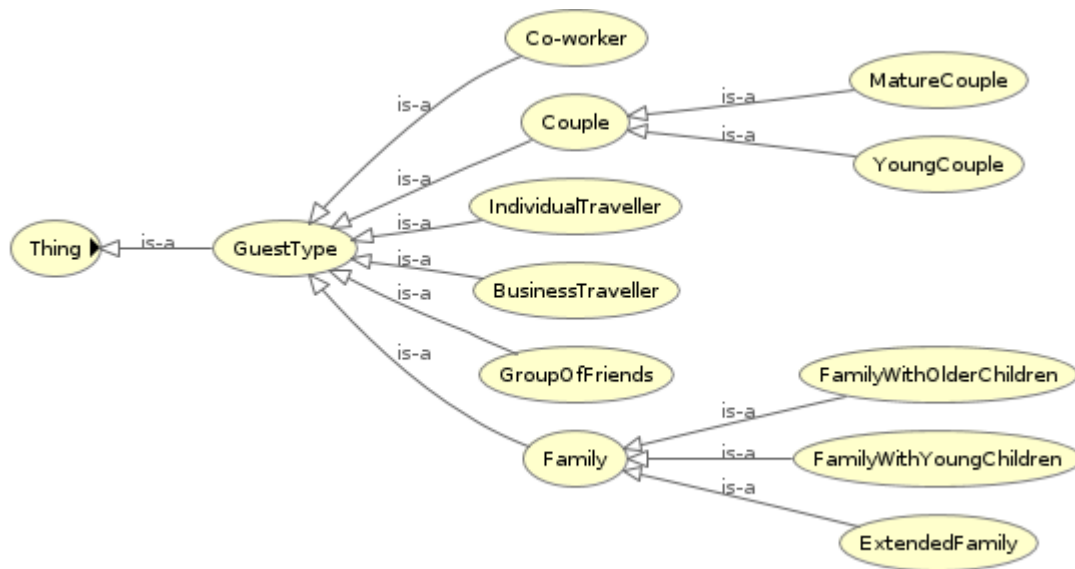
## 4. CONTENT INFORMATION

### 4.1 The natural language(s) of the lexicon

All concepts and properties are defined in Portuguese and some of them in English, Spanish and French. In addition to the main concepts in each language about the accommodation domain, variants whenever they exist were stored.

### 4.2 Entry Type

Considering the section 3.2, the entry types must be from one of the following top-level concepts: *Accommodation*, *Facility*, *Room*, *Service/Staff*, *Guest Type*. We present here the graph representation for *GuestType* concept and his sub-concepts:



#### 4.3 Attributes and their values

Not applicable.

#### 4.4 Coverage of the lexicon

The ontology works on the accommodation domain.

#### 4.5 Intended application of the lexicon

Considering that Hontology can be useful for a wide range of knowledge-based systems, it is useful for a wide range of applications within the accommodation sector, including ontology-based information extraction, text annotation and information visualization.

#### 4.6 POS assignment

Not applicable.

#### 4.7 Reliability (automatically/manually constructed)

The concepts and properties reused from other pre-existent resources were aligned with Hontology which was expanded with new concepts from the content related web sites. These new concepts were filtered out and validated by a user expert.

## 5. RELEVANT REFERENCES AND OTHER INFORMATION

Carvalho, E. and Chaves, M. S. (2012). "Exploring user-generated data visualization in the accommodation sector." In *16th International Conference Information Visualisation*, Montpellier, France, 10-13 July, pp. 198-203. IEEE Computer Society.

Chaves, M. S. and Trojahn, C. (2010). "Towards a multilingual ontology for ontology-driven content mining in social web sites." In *Proceedings of the ISWC 2010 Workshops, Volume I, 1st International Workshop on Cross-Cultural and Cross-Lingual Aspects of the Semantic Web.*, Shanghai, China.

Chaves, M. S., Gomes, R., and Pedron, C. (2012b). "Decision making based on web 2.0 data: The small and medium hotel management." In *Proceedings of the 20<sup>th</sup> European Conference on Information Systems (ECIS)*, Barcelona, Spain, 10-13 June.

Chaves, Marcirio Silveira; Freitas, Larissa A. and Vieira, Renata (2012). "Hontology: a multilingual ontology for the accommodation sector in the tourism industry." In *Proceedings of the 4th International Conference on Knowledge Engineering and Ontology Development*, Barcelona, Spain, 4-7 October.

Pang, B. and Lee, L. (2008). "Opinion Mining and Sentiment Analysis". In *Foundations and Trends in Information Retrieval* 2, 1-2 (Jan. 2008).