Zavod za telekomunikacije

Using Ontologies To Improve Search On WWW.HR Web Directory

Damir Jurić
University of Zagreb
Faculty of Electrical Engineering and Computing
Address: Unska 3, HR-10000 Zagreb, Croatia
E-mail: damir.juric@fer.hr

Introduction

SEMANTIC WEB - vision of World Wide Web in the future where all information's have explicit meaning
ONTOLOGY - definition of rules used for description and representation of some particular field of knowledge

Application that use ontologies seems more “aware” or intelligent from users point of view because they work closer to human conceptual level.

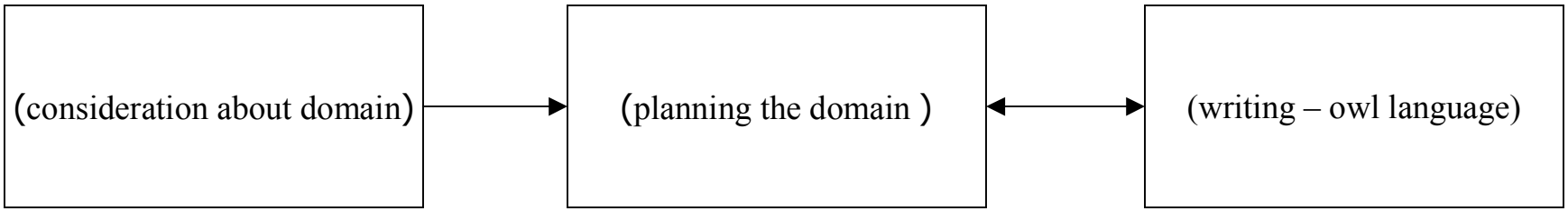
Problem

- to go “beyond” ordinary coincidence of users input with the word in catalog
- to enable searching by the meaning of word
- to create readable structure capable of recording all the data about some concept in the way similar to human level of thinking
- to create application capable of reading, making decisions and presenting data from ontologies

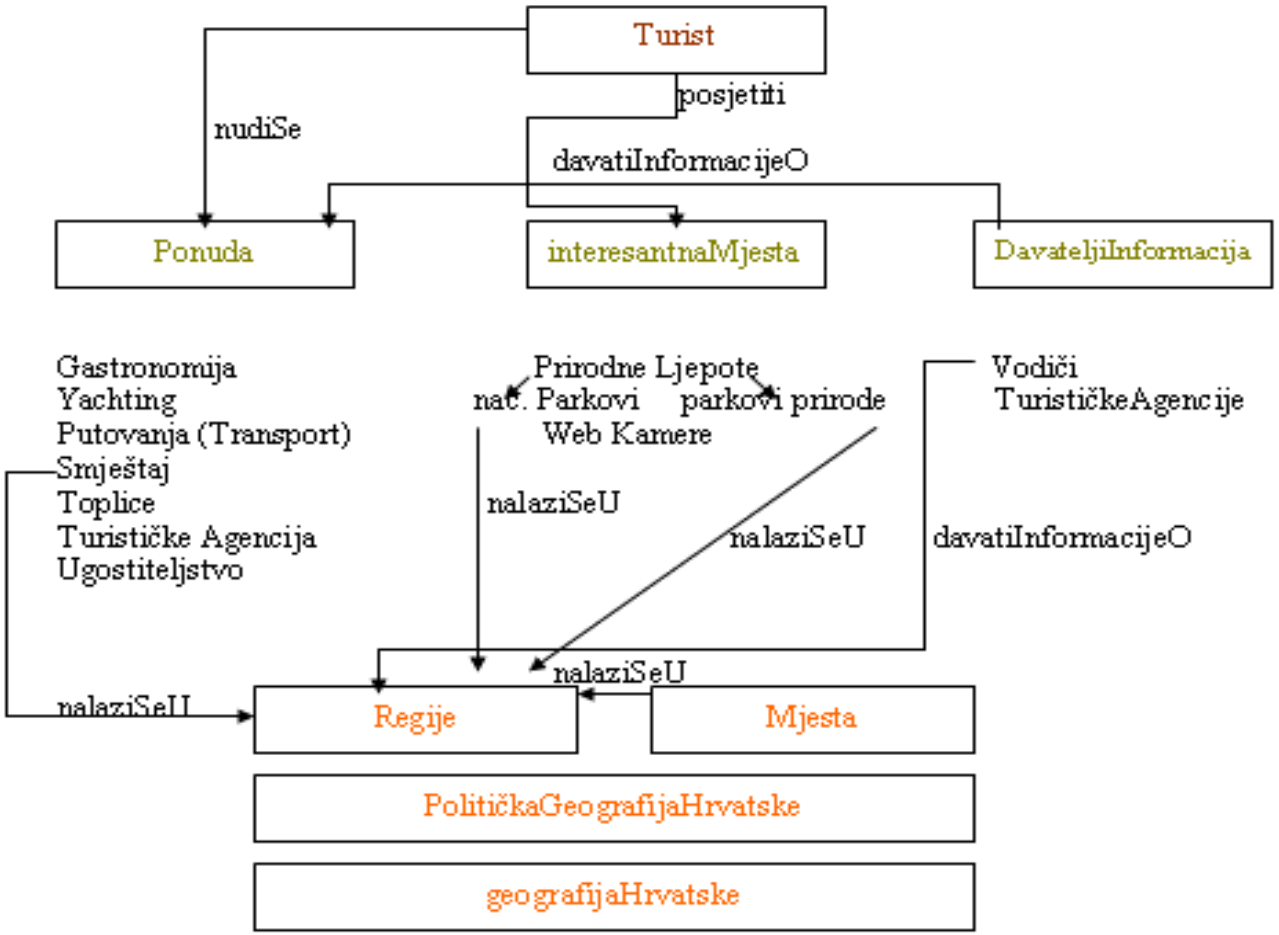
Building ontology

Process of building ontology can be divided in three parts:

- Consideration about domain
- Planning the domain
- Writing the ontology



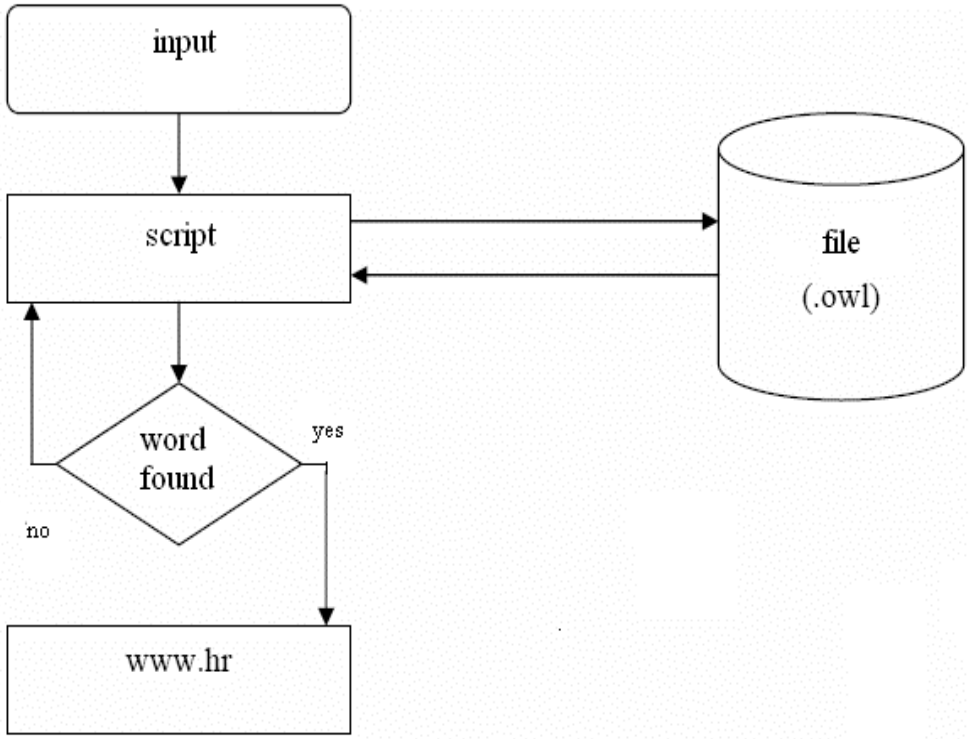
Ontology: turizam.owl




```
<owl:Class rdf:ID="smještaj">
<rdfs:subClassOf
rdf:resource="#ponuda" />
<rdfs:subClassOf
rdf:resource="#turizam" />
<rdfs:subClassOf>
<owl:Restriction>
<owl:onProperty
rdf:resource="#nalaziSeU" />
<owl:allValuesFrom
rdf:resource="#regija" />
</owl:Restriction>
</rdfs:subClassOf>
<rdfs:label>soba</rdfs:label>
<rdfs:label>boravak</rdfs:label>
</owl:Class>
```

Program solution

Program solution is made in the form of PERL script



Things that can be read from ontologies are:

- upclasses
- same classes
- subclasses
- synonyms or alternative words
- restrictions

Unijeli ste pojam:
ponuda

Nadklasa za uneseni pojam je
turizam

Pojam istovrsne klase je:
gastronomija, smještaj, turistička agencija, turistička zajednica, suvenir, vjerski turizam, seoski turizam, kongresni turizam, izlet

Slanje

Podklase za uneseni pojam su:

prehrana
☐ _____ ○ AND ○ OR
zabava
☐ _____ ○ AND ○ OR
restoran
☐ _____ ○ AND ○ OR
yachting
☐ _____ ○ AND ○ OR
putovanja
☐ _____ ○ AND ○ OR
smještaj
☐ _____ ○ AND ○ OR
hotel
☐ _____ ○ AND ○ OR
toplice
☐ _____ ○ AND ○ OR
camping
☐ _____ ○ AND ○ OR
turistička agencija
☐ _____ ○ AND ○ OR
turistička zajednica
☐ _____ ○ AND ○ OR
suvenir
☐ _____ ○ AND ○ OR
ugostiteljstvo
☐

Slanje

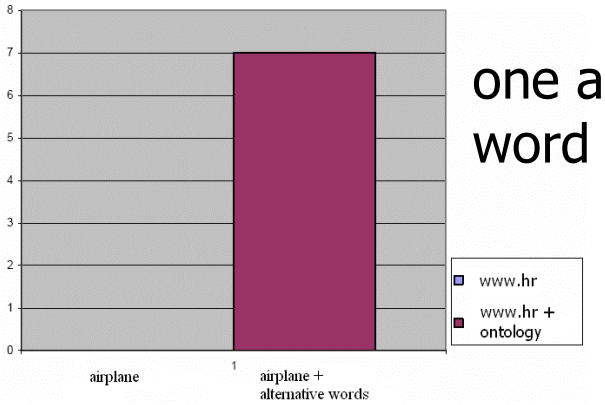
Pretraga za sve podklase!

Output

Results

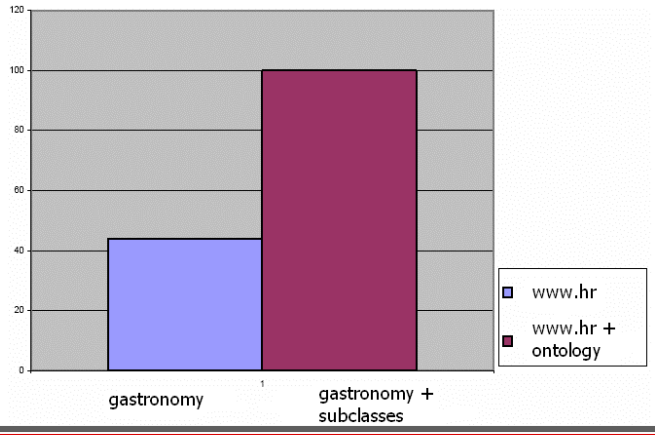
Improved search engine shows considerably improvement when searching with alternative word, subclasses and restrictions

Example 1:
Input = airplane
www.hr: number of found results – 0
www.hr + ontology: found results – 7



one alternative word found

Example 2:
input = gastronomy
www.hr: number of found results – 43
www.hr + ontology: found results – 76



found subclasses: restaurant, tavern, food industry, wine

Results

Example 3:
Input = apartment
apartment + 1 room: found results – 2
apartment + 2 rooms: found results – 20
apartment + 3 rooms: found results – 13
apartment + 4 rooms: found results – 11
apartment + 5 rooms: found results – 1
apartment + kitchen: found results – 9
apartment + terrace: found results – 14
apartment + bathroom: found results – 3
apartment + air conditioning: found results – 1
apartment + living room: found results – 11

Conclusion

- classical searching is already in its full maturity
- although all Web can not be indexed, but only small part of it, number of found result is always huge and more than enough for most of the users
- further standardization of language for writing ontologies is expected, just like development of tools capable of handling with them

References

- [1] World Wide Web Consortium: URL: http://www.w3c.org/2001/sw/WebOnt/
- [2] Roger L. Costello, David B. Jacobs, *OWL Web Ontology Language*, tutorial, The MITRE Corporation, 2003. URL :http://www.xfront.com/
- [3] Đurđica Težak, *Pretraživanje informacija na Internetu: priručnik s vježbama*, Hrvatska sveučilišna naklada, Zagreb. 2002.
- [4] Andrijana Prskalo, *WWW tražilica prilagođena hrvatskom jeziku*, Diplomski rad br. 2186, *Faculty of Electrical Engineering and Computing*, Zagreb, 1997.