
Semantic Web Applications

Prof. Enrico Motta, *PhD*

Knowledge Media Institute
The Open University
Milton Keynes, UK



The Semantic Web

The collection of all formal, machine processable, web accessible, ontology-based statements (semantic metadata) about web resources and other entities in the world, expressed in a knowledge representation language based on an XML syntax (e.g., OWL, DAML, DAML+OIL, RDF, etc...)

KMi

Search

Home News Projects Technologies Publications People

People [People/All Members]

Members [76] [A][B][C][D][E][F][G][H][I][J][K][L][M][N][O][P][Q][R][S][T][U][V][W][X][Y][Z]

Affiliates [5]

Alumni [43]

Paul Alexander
IT & Systems Manager [info] [homepage] [email] [RDF/XML]

Carlo Allocca
Intern [info] [email] [RDF/XML]

Sofia Angeletou
Research Student [info] [email] [RDF/XML]

Michelle Suzanne Bachler
Project Officer [info] [homepage] [email] [RDF/XML]

David Bainbridge
Visiting Researcher [info] [homepage] [RDF/XML]

Claudio Baldassarre
Research Assistant [info] [homepage] [email] [RDF/XML]

KMi

Search

Home News Projects Technologies Publications People

People [People/All Members]

Members [76] [A][B][C][D][E][F][G][H][I][J][K][L][M][N][O][P][Q][R][S][T][U][V][W][X][Y][Z]

Affiliates [5]

Alumni [43]

Paul Alexander
IT & Systems Manager [info] [homepage] [email] [RDF/XML]

Carlo Allocca
Intern [info] [email] [RDF/XML]

Sofia Angeletou
Research Student [info] [email] [RDF/XML]

Michelle Suzanne Ba
Project Officer [info] [RDF/XML]

David Bainbridge
Visiting Researcher [info] [RDF/XML]

Claudio Baldassarre
Research Assistant [info] [RDF/XML]

Semantic Web Document

```
<foaf:PersonalProfileDocument rdf:about="http://kmi.open.ac.uk/people/sofia-angeletou" >
  <dc:title>Sofia Angeletou's RDF Description</dc:title>
  <rdfls:label>Sofia Angeletou's RDF Description</rdfls:label>
  <dc:description>RDF description for Sofia Angeletou in machine-readable form</dc:description>
  <dc:creator rdf:resource="http://identifiers.kmi.open.ac.uk/people/sofia-angeletou">
  <foaf:maker rdf:resource="http://identifiers.kmi.open.ac.uk/people/sofia-angeletou">
  <foaf:primaryTopic rdf:resource="http://identifiers.kmi.open.ac.uk/people/sofia-angeletou">
</foaf:PersonalProfileDocument>
```

Semantic Content is Increasing....



DBpedia.org

Querying Wikipedia like a Database.

Enrico Motta
Professor of Knowledge Technologies [info] [homepage] [...]

Liza Mu
Visiting Researcher [info] [email] [RDF/XML]



```
<rdf:Description rdf:about="">
  <mediapro:People>
    <rdf:Bag>
      <rdf:li>Jim Hendler</rdf:li>
      <rdf:li>Enrico</rdf:li>
    </rdf:Bag>
  </mediapro:People>
</rdf:Description>
```

AKT Triplestore Browser

about browse manage query demo

Southampton ECS People
file:///usr/local/share/akt/Southampton/southampton-people.rdf

creator [AKT Project](#)
[Nick Gibbins](#)

description This ontology contains information about the members of the Department of Electronics and Computer Science at the University of Southampton.

title Southampton ECS People

created 2004-02-04

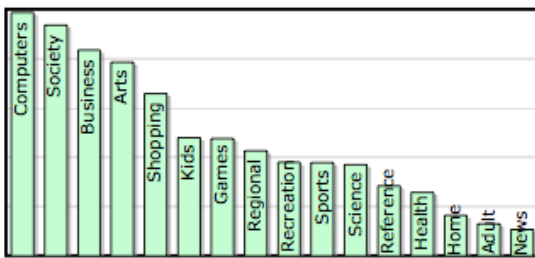
label Southampton ECS People

Shenley Church End ca. 76 m
[Shenley](#)
United Kingdom
populated place
N 52° 1' 0" W 0° 47' 0"
52.01667 / -0.78333

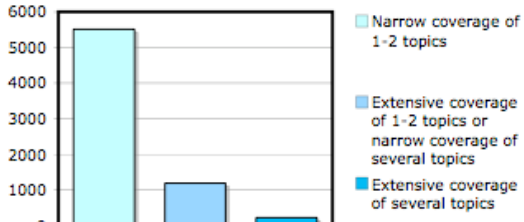
zoom move edit history tag delete alternate names
perma link semantic web rdf

```
<rdf:RDF>
  <Feature rdf:about="http://sws.geonames.org/5201667">
    <name>Shenley Church End</name>
    <alternateName>Shenley</alternateName>
    <inCountry>United Kingdom</inCountry>
    <rdf:resource="http://www.geonames.org/counties">
      </rdf:RDF>
```

Domain Coverage on the SW



Distribution of documents in the 16 top categories of DMOZ



Distribution of the documents according to their coverage

- Great variety: Some topics are almost not covered (e.g. Health), while some are covered reasonably well (e.g. Society, Computers)
- As we can expect, a large number of narrow coverage documents and a small number of large coverage ones.

The Rise of Semantics



ORACLE
TECHNOLOGY NETWORK

PRODUCTS
Database
Middleware
Developer Tools
Enterprise Management
Applications Technology
Extensions and Plugins
Products A-Z

TECHNOLOGIES
BI & Data Warehousing
Java
Linux
.NET
Office
PHP
Security
Service-Oriented Architecture
...

Getting Started | Downloads | Documentation | Forums

Semantic Technologies Center

Semantic Technologies are designed to extend the capabilities of inf enterprise systems to be networked in meaningful ways. The adoptor Consortium (W3C) standards like XML, RDF (Resource Description F Ontology Language) serve as foundation technologies to advancing t technologies.

Oracle Spatial 10g introduces the industry's first open, scalable, secure and reliable RDF ma graph data model. RDF triples are persisted, indexed and queried, similar to other object-rel; 10g RDF database ensures that application developers benefit from the stability of Oracle's secure semantic applications. Application areas include:

- Life Sciences: Biological pathway analysis, discovery an
- Defense & Intelligence: Data and content integration, rea
- Enterprise Application Integration: Data and systems inte services
- CRM/ERP: Supply chain integration, sourcing optimizati

Gartner

Finding and Exploiting Value in Semantic Technologies on the Web
9 May 2007
David W. Cearley | Whit Andrews | Nicholas Gall

garlik
Home
About Garlik
DataPatrol
Advice
Sign up

powerfulstuff
sign in

Revolutionary online
DataPatrol
monitoring and monitor information
Home | Products | Services & industry solutions | Support & downloads | My IBM

DeveloperWorks
Web development | Open source | XML

The future of the Web is Semantic
Ontologies form the backbone of a whole new way to understand online data
Level: Introductory

Key Points



- The SW is less and less an aspiration and more and more a reality: there is already a large amount of semantic markup available on the web
- However the domain coverage of the semantic web is patchy. Some domains are well covered, others covered very little
- This emerging large scale semantics opens up new possibilities for building novel, knowledge-based applications
- In addition, it may also provide a solution to one of the holy grails of AI research: the availability of large-scale background knowledge to enable intelligent behaviour



Semantic Web Applications

Application Types



- Lots of different types...diversity is increasing...see “Intro to SW” talk.....
 - Enterprise Information Integration
 - Web 2.0 like
 - Devices
 - Agents
 - Etc...
- Here we focus on two broad classes of applications
 - (Semantically) Closed SW Applications (*1st generation*)
 - Typically in the domain of Enterprise Information Integration
 - (Semantically) Open SW Applications (*2nd generation*)



Closed SW Applications

AKT CS AKTive Space Take a tour

About this page research area/region region/research area

Research area **Radial:** 100 miles **Map:** uk-political

Researcher
Top 5 10 20 unlimited
Order by Grant total RAE result

Information Systems
information interfaces and presentation
information systems applications
information storage and retrieval
database management
general

Computing Methodologies
document and text processing
simulation and modeling
pattern recognition
image processing and computer vision
computer graphics

artificial intelligence
symbolic and algebraic manipulation
general

Computer Applications
computers in other systems
computer-aided engineering
arts and humanities
social and behavioral sciences
life and medical sciences

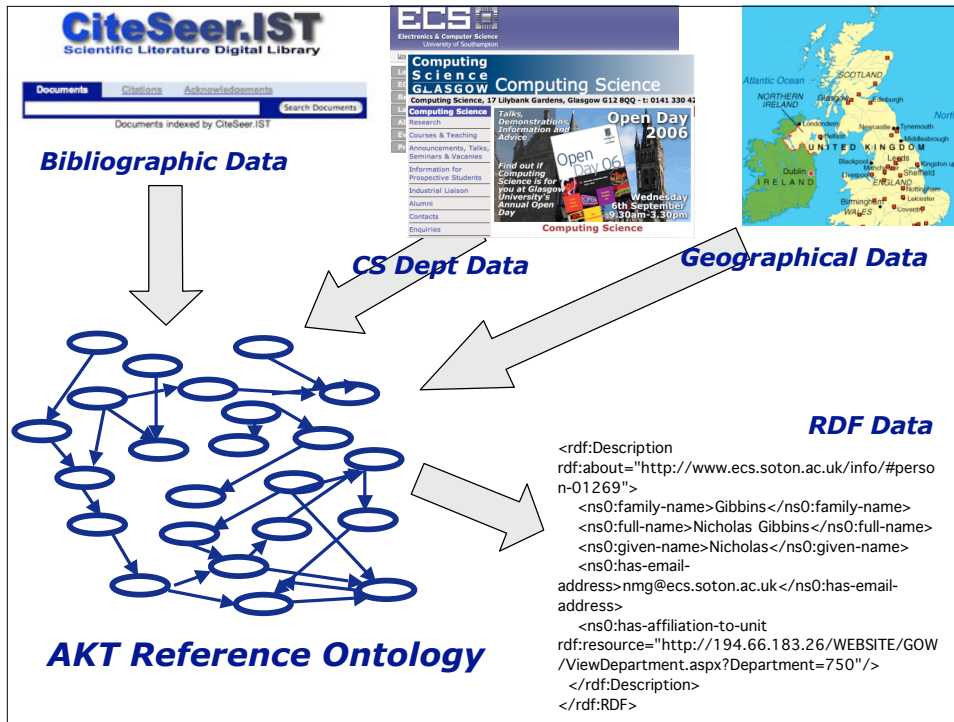
NR Shadbolt
LA Carr
DC De Roure
NR Jennings
L Moreau

Overview: NR Shadbolt
[browse](#)

Name NR Shadbolt
Institution Intelligence, Agents and Multimedia, University of Southampton
Email nrs@ecs.soton.ac.uk
Tel +442380597682
Fax +442380592865

CoP mkw
E Motta
Kieron
OHara
W Hall
A Tate
Ian Millard
Les Carr
Y Wilks

Research Networks and Distributed Systems



AKT CS AKTive Space Take a tour

About this page research area/region region/research area

Research area: Radial: 100 miles Map: uk-political

Research area

- information systems and presentation
- information systems applications
- information storage and retrieval
- database management
- general
- Computing Methodologies**
- document and text processing
- simulation and modeling
- pattern recognition
- image processing and computer vision
- computer graphics
- artificial intelligence**
- symbolic and algebraic manipulation
- general
- Computer Applications**
- computers in other systems
- computer-aided engineering
- arts and humanities
- social and behavioral sciences
- life and medical sciences

Researcher

Top 5 10 20 unlimited

Order by Grant total RAE result

- NR Shadbolt
- LA Carr
- DC De Roure
- NR Jennings
- L Moreau

Overview: NR Shadbolt

[browse](#)

Name NR Shadbolt
Institution Intelligence, Agents and Multimedia, University of Southampton
Email nrs@ecs.soton.ac.uk
Tel +442380597682
Fax +442380592865

Fluid Dynamics
Aerodynamics
Design and Testing Technology
Biological Sciences Domain
Image and Vision Computing
Networks and Distributed Systems

CoP mkw
E Motta
Kieron
OHara
W Hall
A Tate
Ian Millard
Les Carr
Y Wilks

FLINK
The Who is Who of the Semantic Web

home network ontology cluster world stats feedback about

Social Network of Enrico Motta

Statistics

Indegree	67.0 (67.0) -
Closeness	0.5 (0.5) -
Betweenness	8212.68 (8212.68) -
Top Publications	53.0 (53.0) -
Impact	11.48 (11.48) -

What do the numbers mean?

Rankings

Indegree	11 (11) -
Closeness	155 (155) -
Betweenness	10 (10) -
Top Publications	13 (13) -
Impact	193 (193) -

Links

- Search Google!
- Search Foafnaut!
- Search Citeseer!
- Search A9 (Amazon)!

Got FOAF?

Tip: You can drag nodes with your mouse. You can also use the scrollbars to pan the image. You can zoom in and out using the +/- buttons, but also with the roller.

Closed SW Applications

- A 'corporate ontology' is used to provide a homogeneous view over heterogeneous data sources.
- Limited use of existing SW data
 - Data are typically scraped from existing non-semantic sources
- Typically closed to additional (heterogenous) semantic resources
 - On-the-fly semantic data integration is usually not supported
- Often tackle Enterprise Information Integration scenarios
- Can be seen (to some extent) as the application of traditional knowledge engineering technologies in a web environment

garlik
powerful stuff

[About us](#) [DataPatrol Advanced](#) [Advice](#) [Sign up](#) [Log in](#)

NEW DataPatrol Advanced - the best way to protect your identity online

DataPatrol Advanced shows you the information that's stored about you online, from public records to entries on websites. Together with regular risk assessments and our expert advice, it helps you protect yourself from identity theft.

- Sign up now for £2.99 per month and get the first month free
- Learn more

Revolutionary online protection for £2.99 per month - first month FREE

DataPatrol Advanced is your first line of defence against identity theft, including:

- FREE** Credit report
- FREE** Local area report
- FREE** Digital identity check

➤ Sign up for DataPatrol Advanced now

Who are garlik?

We're the founders of internet banks Egg and First Direct, leading technology innovators and privacy experts who've teamed up to protect you online. We want to give you and your family more power over the way your personal information is used in the digital world.

➤ Find out more about garlik

How to keep yourself safe online

Are you worried about identity theft? In the UK, we're each worth an average of £85,000 to an online fraudster. We don't want you to be scared. But we do want you to be safe. Our experts separate fact from fiction and help you keep your personal information secure.

➤ Sign up for DataPatrol Advanced now

Existing users: Login/Upgrade

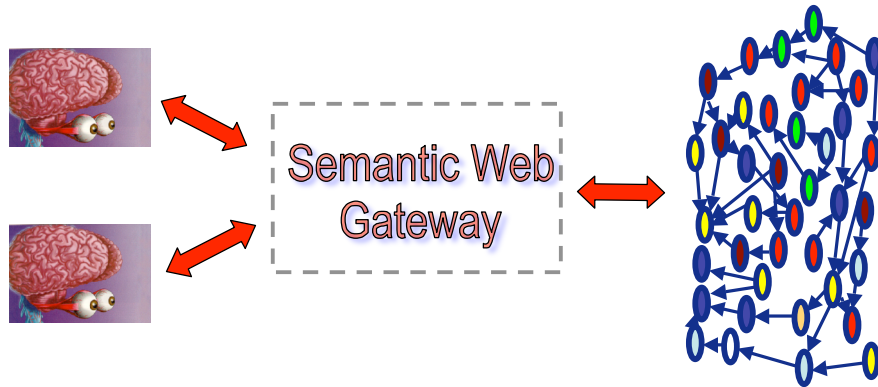
E-mail: ➤ Go

Password:

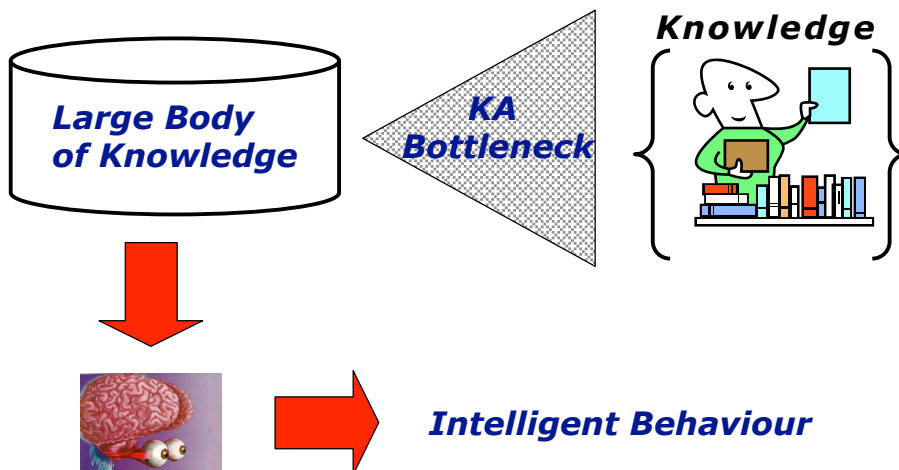
Open SW Applications

(2nd Generation SW Applications)

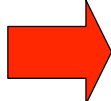
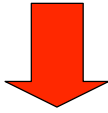
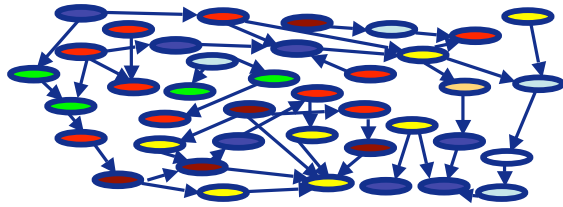
Architecture of OSW Apps



The Knowledge Acquisition Bottleneck

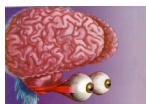


SW as Enabler of Intelligent Behaviour



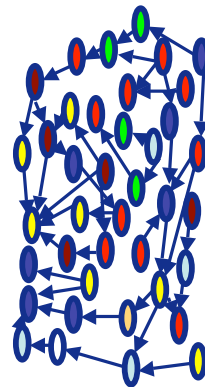
Intelligent Behaviour

Swoogle



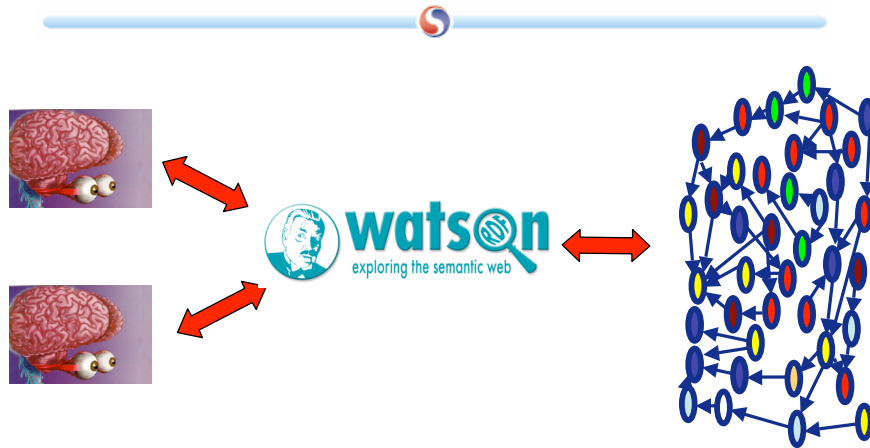
Swoogle

semantic web search 2006



<http://swoogle.umbc.edu/>

Watson

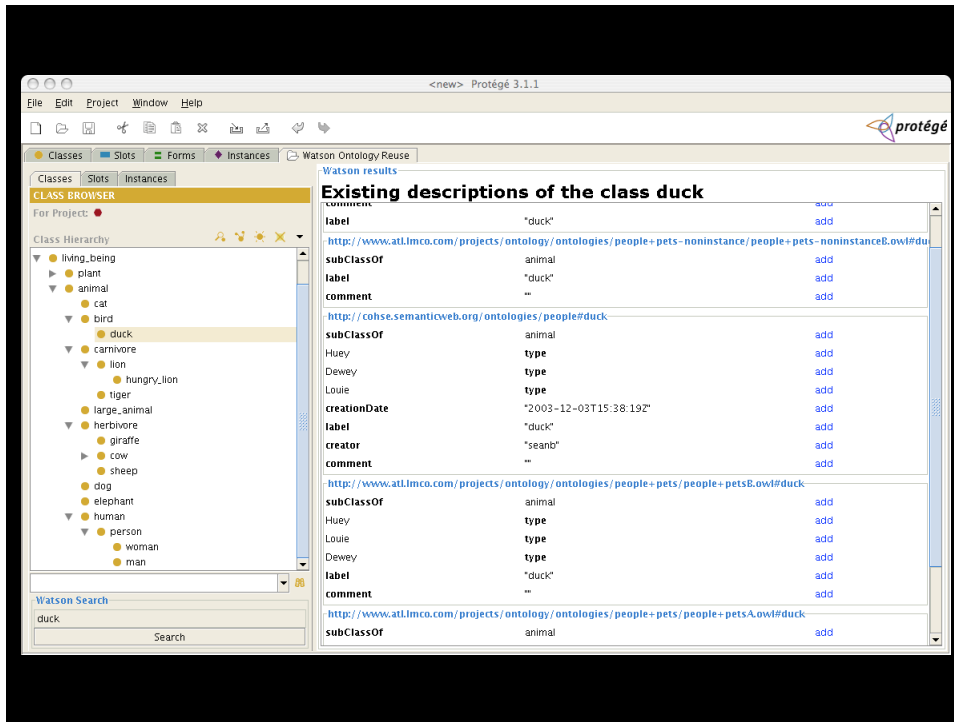



<http://watson.kmi.open.ac.uk>



watson
exploring the semantic web

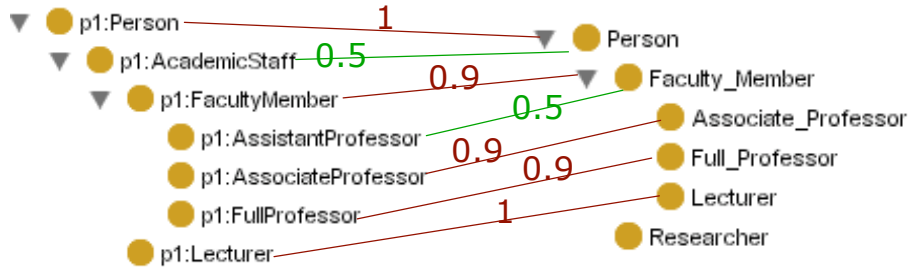
- Sophisticated quality control mechanism
 - Detects duplications
 - Fixes obvious syntax problems
 - E.g., duplicated ontology IDs, namespaces, etc..
- Structures ontologies in a network
 - Using relations such as: *extends*, *inconsistentWith*, *duplicates*
- Provides efficient API
- Supports formal queries (SPARQL)
- Variety of ontology ranking mechanisms
- Modularization/Combination support
- Plug-ins for Protégé and NeOn Toolkit
- Very cool logo!





Example: Using the SW as Background Knowledge for Ontology Matching

Ontology Matching



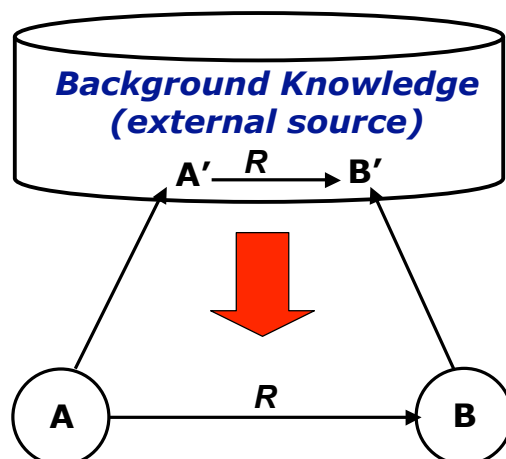
-Label similarity methods

- e.g., Full_Professor = FullProfessor

-Structure similarity methods

- Using taxonomic/property related information

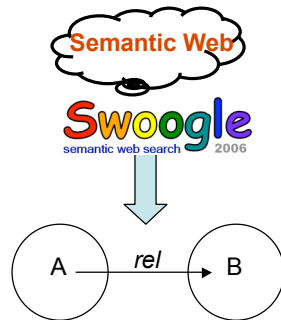
New paradigm: Use of Background Knowledge



External Source = SW

Proposal:

- rely on online ontologies (Semantic Web) to derive mappings
- ontologies are **dynamically** discovered and combined

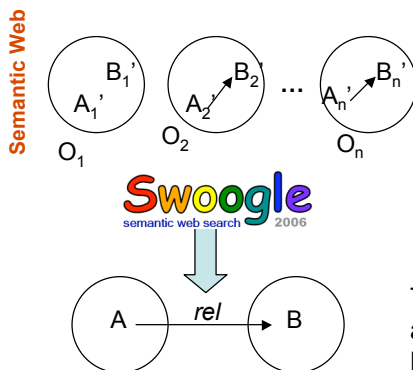


Does not rely on any pre-selected knowledge sources.

M. Sabou, M. d'Aquin, E. Motta, "Using the Semantic Web as Background Knowledge in Ontology Mapping", Ontology Mapping Workshop, ISWC'06. **Best Paper Award**

Strategy 1 - Definition

Find ontologies that contain equivalent classes for A and B and use their relationship in the ontologies to derive the mapping.



For each ontology use these rules:

$$A' \equiv B' \Rightarrow A \equiv B$$

$$A' \subseteq B' \Rightarrow A \subseteq B$$

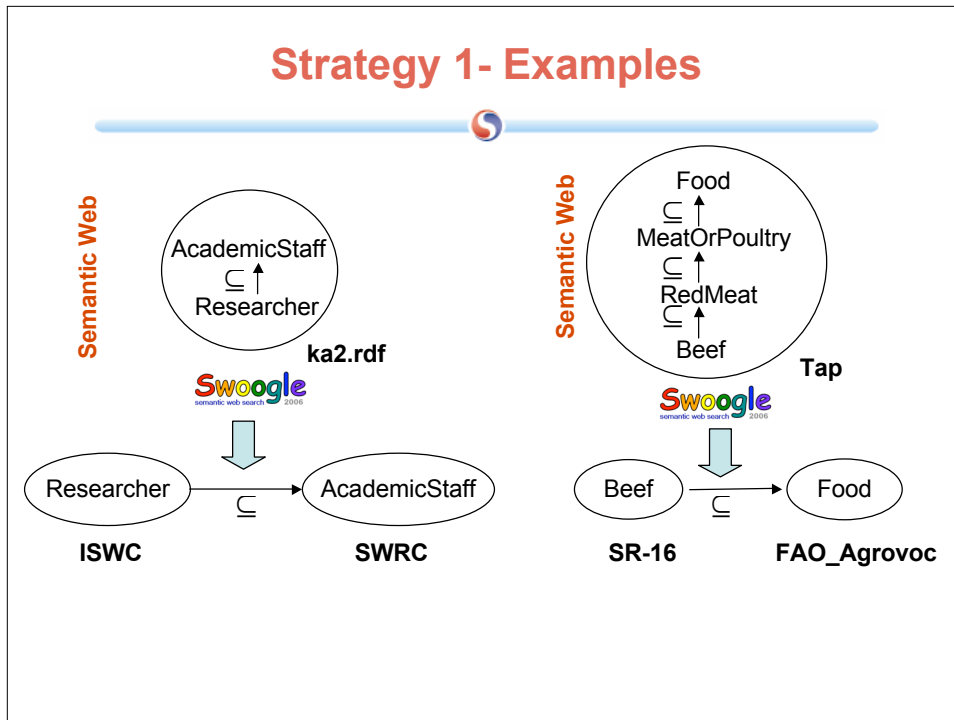
$$A' \supseteq B' \Rightarrow A \supseteq B$$

$$A' \perp B' \Rightarrow A \perp B$$

These rules can be extended to take into account indirect relations between A' and B', e.g., between parents of A' and B':

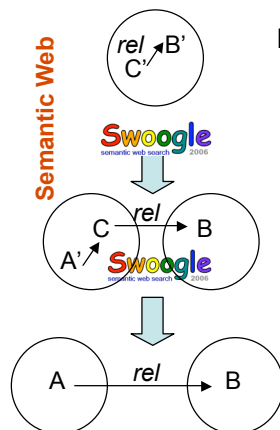
$$A' \subseteq C \wedge C \perp B' \Rightarrow A' \perp B'$$

Strategy 1- Examples



Strategy 2 - Definition

Principle: If no ontologies are found that contain the two terms then combine information from multiple ontologies to find a mapping.



Details:

- (1) Select all ontologies containing A' equiv. with A
- (2) For each ontology containing A' :

(a) if $A' \subseteq C$ find relation between C and B .

(b) if $A' \supseteq C$ find relation between C and B .

$$(r1) A' \subseteq C \wedge C \subseteq B \Rightarrow A \subseteq B$$

$$(r2) A' \subseteq C \wedge C \equiv B \Rightarrow A \equiv B$$

$$(r3) A' \subseteq C \wedge C \perp B \Rightarrow A \perp B$$

$$(r4) A' \supseteq C \wedge C \supseteq B \Rightarrow A \supseteq B$$

$$(r5) A' \supseteq C \wedge C \equiv B \Rightarrow A \supseteq B$$

Strategy 2 - Examples

Ex1: Chicken Vs. Food

$$\left. \begin{array}{l} \text{Chicken} \subseteq \text{Poultry} \text{ (midlevel-onto)} \\ \text{Poultry} \subseteq \text{Food} \text{ (Tap)} \end{array} \right\} \xrightarrow{(r1)} \text{Chicken} \subseteq \text{Food}$$

(Same results for Duck, Goose, Turkey)

Ex2: Ham Vs. Food

$$\left. \begin{array}{l} \text{Ham} \subseteq \text{Meat} \text{ (pizza-to-go)} \\ \text{Meat} \subseteq \text{Food} \text{ (SUMO)} \end{array} \right\} \xrightarrow{(r1)} \text{Ham} \subseteq \text{Food}$$

Ex3: Ham Vs. Seafood

$$\left. \begin{array}{l} \text{Ham} \subseteq \text{Meat} \text{ (pizza-to-go)} \\ \text{Meat} \perp \text{Seafood} \text{ (wine.owl)} \end{array} \right\} \xrightarrow{(r3)} \text{Ham} \perp \text{Seafood}$$

Large Scale Evaluation

Matching AGROVOC (16k terms) and NALT(41k terms)

	Nr.	Examples
Subclass ($\xrightarrow{\subseteq}$)	1477	$Lamb \xrightarrow{\subseteq} Sheep, Soap \xrightarrow{\subseteq} Detergent, Asbestos \xrightarrow{\subseteq} Pollutant$ $Oasis \xrightarrow{\subseteq} Ecosystem, RAM \xrightarrow{\subseteq} ComputerEquipment$
SuperClass ($\xrightarrow{\supseteq}$)	1857	$Shop \xrightarrow{\supseteq} Supermarket, Spice \xrightarrow{\supseteq} BlackPepper, Valley \xrightarrow{\supseteq} Canyon$ $Infrastructure \xrightarrow{\supseteq} Highway, Storm \xrightarrow{\supseteq} Tornado, Rock \xrightarrow{\supseteq} Crystal$
Disjoint ($\xrightarrow{\perp}$)	229	$Fluid \xrightarrow{\perp} Solid, Fluid \xrightarrow{\perp} Gas, Pond \xrightarrow{\perp} River, Plant \xrightarrow{\perp} Animal$ $Newspaper \xrightarrow{\perp} Journal, Fruit \xrightarrow{\perp} Vegetable, Female \xrightarrow{\perp} Male$
Total	3563	

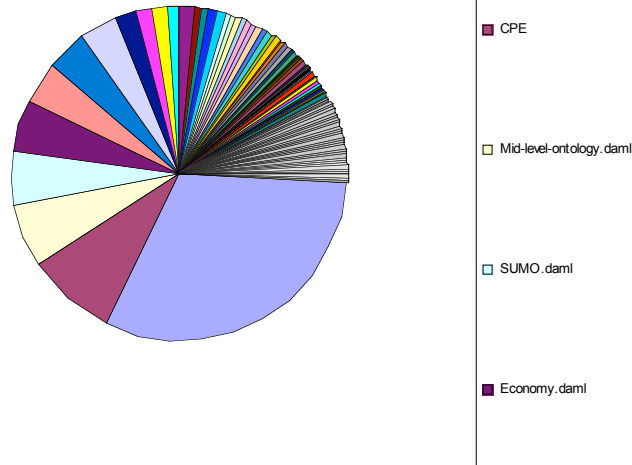
(derived from 180 different ontologies)

Evaluation: 1600 mappings, two teams, 70% Precision

M. Sabou, M. d'Aquin, E. Motta. "Using the Semantic Web as Background Knowledge for Ontology Matching". To appear in the Journal of Data Semantics

Chart 2

Ontologies (180) used to derive mappings.



Error Type	Nr./ %	Examples				
		AGROVOC Concept	Labels	Rel.	NALT Concept	Labels
Anchor	114, 53%	<i>c.6443</i>	Rams, Tups	\sqsubseteq	<i>memory</i>	memory
		$O_1:ram \sqsubseteq O_1:memory$ $O_1 = \text{http://www.arches.uga.edu/~gonen/qos_bilal.owl}$				
Subsumption as generic relation	40, 18%	<i>c.3954</i>	Irrigation	\sqsubseteq	<i>agriculture</i>	agriculture
		$O_1:Irrigation \sqsubseteq O_1:SoilCultivation \sqsubseteq O_1:Agriculture$ $O_1 = \text{http://sweet.jpl.nasa.gov/ontology/human_activities.owl}$				
Subsumption as part-whole	16, 7%	<i>c.666</i>	Asia	\sqsupseteq	<i>Iran</i>	Iran
		$O_1:Asia \sqsupseteq O_1:WestAsia \sqsupseteq O_1:Iran$ $O_1 = \text{http://islab.hanyang.ac.kr/damls/Country.daml}$				
Subsumption as role	11, 5%	<i>c.11091</i>	Garlic	\sqsubseteq	<i>ingredients</i>	ingredients
		$O_1:garlic \sqsubseteq O_1:vegetable \sqsubseteq O_1:ingredient$ $O_1 = \text{http://cvs.sourceforge.net/viewcvs.py/instancestore/instancestore/ontologies/Attic/pizza9.daml?rev=1.2}$				
Inaccurate labeling	12, 5%	<i>c.1693</i>	Coal	\sqsubseteq	<i>industry</i>	industry
		$O_1:coal \sqsubseteq O_1:industry$ $O_1 = \text{http://www.aifb.uni-karlsruhe.de/WBS/meh/mapping/data/russia1a.rdf}$				
Different View	12, 5%	<i>c.2943</i>	Fishes	\sqsupseteq	<i>lobsters</i>	lobsters
		$O_1:Fish \sqsupseteq O_1:MarineInvertebrate \sqsupseteq O_1:Crustacean \sqsupseteq O_1:Lobster$ $O_1 = \text{http://139.91.183.30:9090/RDF/VRP/Examples/tap.rdf}$				

Conclusions



- Many 'classic' applications of SW technology use ontologies to support the integration of distributed data sources
 - These applications are typically 'semantically closed'
- As more and more semantic information becomes available on the SW, researchers are also looking at 'semantically open' applications, able to exploit large scale semantics to support intelligent problem solving
- This approach is being used in a number of other scenarios, including:
 - Semantic Web Browsing
 - Question Answering
 - Integration of Folksonomies with the SW

Selected References



- Ontology Matching
 - Lopez, V., Sabou, M., Motta, E. (2006). *"Mapping the real semantic web on the fly"*. ISWC 2006
 - Sabou, M., D'Aquin, M., Motta, E. (2006). *"Using the semantic web as background knowledge for ontology mapping"*. ISWC 2006 Workshop on Ontology Mapping.
- Integration of Web2.0 and Semantic Web
 - L.Specia, E. Motta, *"Integrating Folksonomies with the Semantic Web"*, ESWC 2007.
 - Angeletou, S., Sabou, M., Specia, L., and Motta, E., (2007). *"Bridging the Gap Between Folksonomies and the Semantic Web: An Experience Report"*. ESWC 2007 Workshop on Bridging the Gap between Semantic Web and Web 2.0.
- Watson
 - d'Aquin, M., Sabou, M., Dzbor, M., Baldassarre, C., Gridinoc, L., Angeletou, S. and Motta, E.: *"WATSON: A Gateway for the Semantic Web"*. Poster Session at ESWC 2007

'Vision' Papers



- Motta, E., Sabou, M. (2006). **"Next Generation Semantic Web Applications"**. 1st Asian Semantic Web Conference, Beijing.
- Motta, E., Sabou, M. (2006). **"Language Technologies and the Evolution of the Semantic Web"**. LREC 2006, Genoa, Italy.
- Motta, E. (2006). **"Knowledge Publishing and Access on the Semantic Web: A Socio-Technological Analysis"**. IEEE Intelligent Systems, Vol.21, 3, (88-90).

KNOWLEDGE MEDIA

KMi
I N S T I T U T E