



# **What does it mean to be semantic?** **(On the role of semantics in the Semantic Web)**

**Enrico Motta**  
**Knowledge Media Institute**  
**The Open University**  
**United Kingdom**

# Main objective of the talk

- To reflect on the nature of Semantic Web applications and discuss questions such as:
  - What does it mean to do semantics on the web? – i.e., what are the **innovative** elements that distinguish Semantic Web applications from other systems, in particular conventional DBs and KBS?
  - What makes Semantic Web applications interesting and powerful –i.e., in what ways the use of web semantics provides us with a competitive advantage over other approaches?



How does semantics make us stronger?

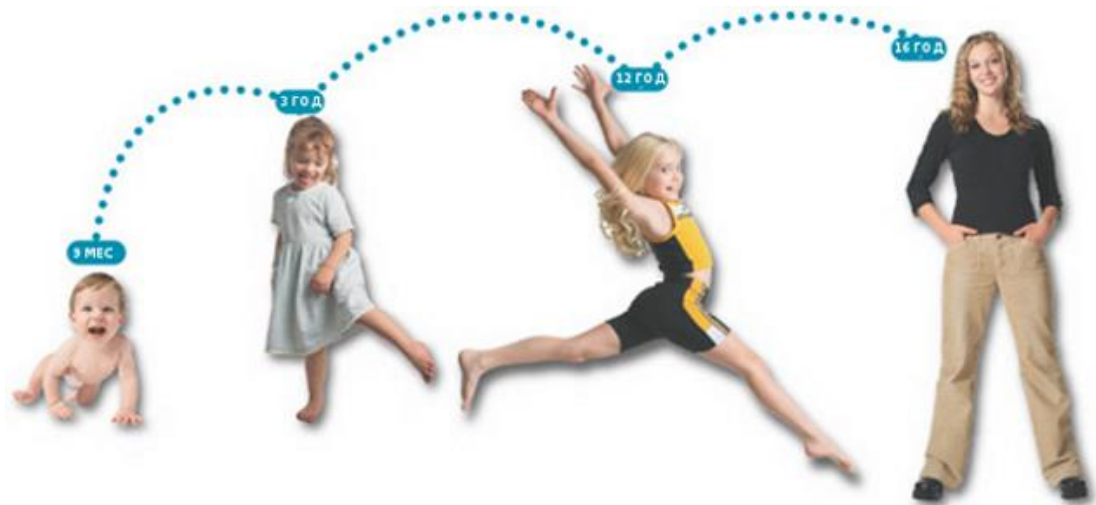
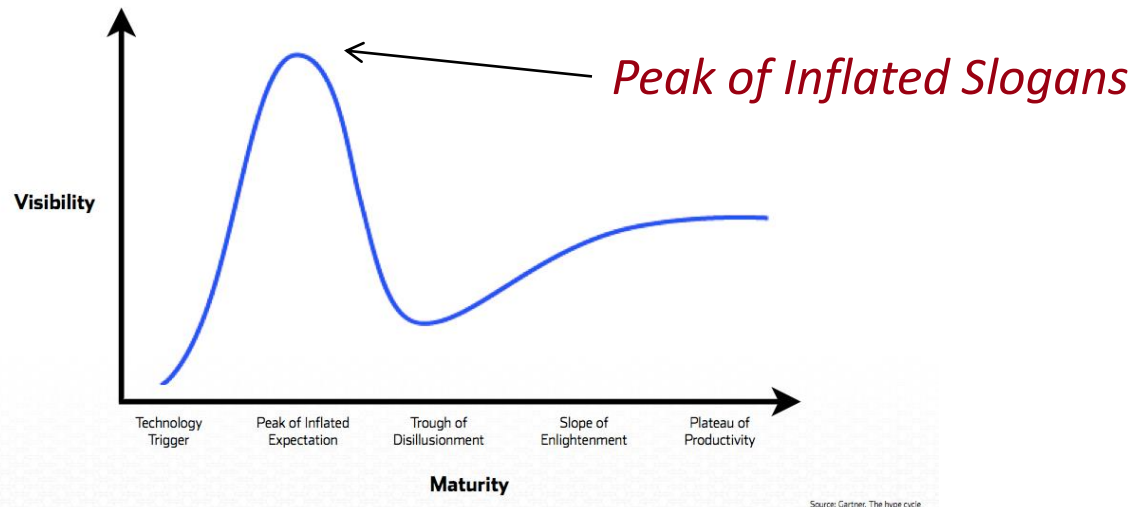


# Motivation

Why should we reflect on these issues?

# The SW as a teenager: time to reflect

- For the past 14 years we have been busy nurturing and protecting our child and telling everybody that she was going to change the world
- The child is now entering her teens, is reasonably healthy, and we are no longer so worried about her survival. Hence, it is time to worry instead about whether she is growing beautiful and intelligent....
- Also because she may look a bit different from what at least some of us had anticipated....



# The original vision



May 17, 2001

## The Semantic Web

**A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities**

By Tim Berners-Lee, James Hendler and Ora Lassila

The entertainment system was belting out the Beatles' [REDACTED] when the phone rang. When Pete answered, his phone turned the sound down by sending a message to all the other *local* devices that had a *volume control*. His sister, Lucy, was on the line from the doctor's office: "Mom needs to see a specialist and then has to have a series of physical therapy sessions. Biweekly or something. I'm going to have my agent set up the appointments." Pete immediately agreed to share the chauffeuring.

At the doctor's office, Lucy instructed her Semantic Web agent through her handheld Web browser. The agent promptly retrieved information about Mom's *prescribed treatment* from the doctor's agent, looked up several lists of *providers*, and checked for the ones *in-plan* for Mom's insurance within a *20-mile radius* of her *home* and with a *rating of excellent or very good* on trusted rating services. It then began trying to find a match between available *appointment times* (supplied by the agents of individual providers through their Web sites) and Pete's and Lucy's busy schedules. (The emphasized keywords indicate terms whose semantics, or meaning, were defined for the agent through the Semantic Web.)

In a few minutes the agent presented them with a plan. Pete didn't like it—University Hospital was all the way across town from Mom's place, and he'd be driving back in the middle of rush hour. He set his own agent to redo the search with stricter preferences about *location* and *time*. Lucy's agent, having *complete trust* in Pete's agent in the context of the present task, automatically assisted by supplying access certificates and shortcuts to the data it had already sorted through.



A Letter from the Editor



**Where Are All the  
Intelligent Agents?**

# What's the big deal?

One could wonder what's the problem. Why is this an issue? We know what SW applications are about...

They are about using OWL/RDF

They are about ontologies

They are about linked data

Etc..

Let's look at a couple of concrete examples from the last ISWC...

# Rexplore

Making Sense of Scholarly Data

Info Migrations

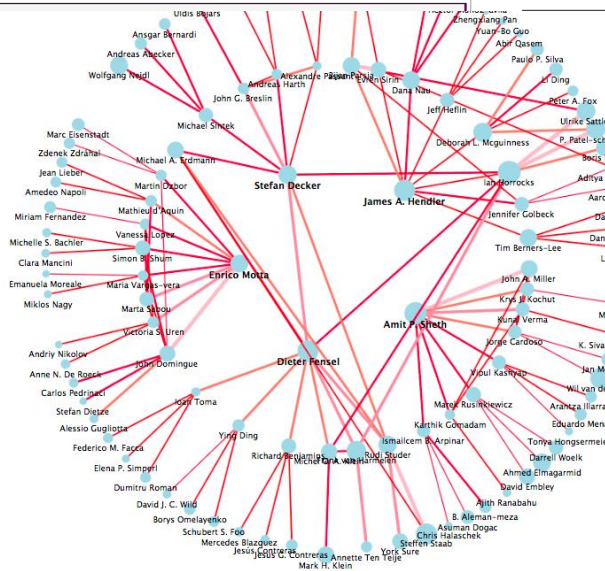
## Semantic Web

Publications: 22 143  
Citations: 120 704

- Plot authors and publications
- Plot average citations vs authors and publications
- Explore authors
- Explore publications

## Semantic Relationships

- Reload Semantic Web relationships
- Deactivate semantic layer

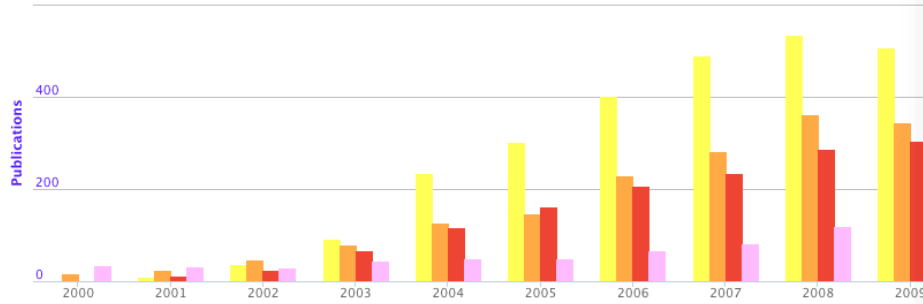


## Semantic Web

broaderGeneric Topics - publication trend

Semantic Web Service Semantic Annotation Semantic Web Technology Linked Data

Zoom 5y 10y 15y 20y 30y All



World Wide Web

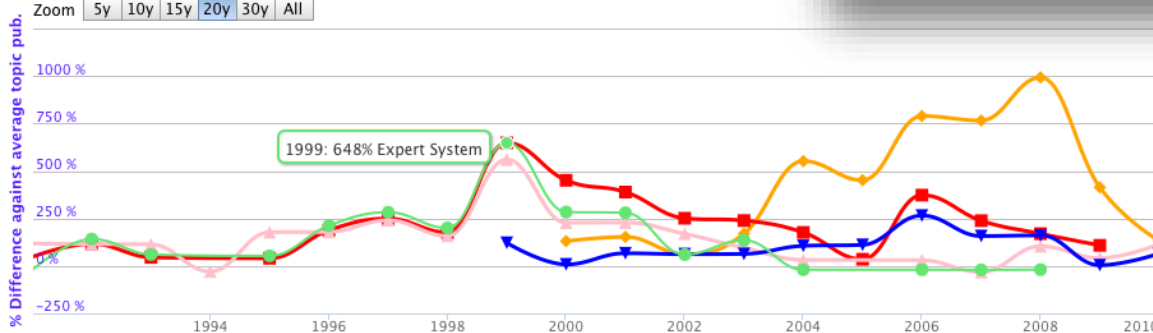
- cT - Semantic Web -
- bG - Semantic Web Technology -
- bG - Semantic Web Rule Language
- bG - Web of Data
- cT - Semantic Technologies +
- cT - Semantic Search
- cT - Semantic Metadata
- cT - Social Web
- cT - Linked Open Data
- bG - Semantic Web Service +
- bG - Semantic Annotation -
- bG - Semantic Metadata
- bG - Semantic Wiki

## Enrico Motta

Publications normalized by topics

total publications Semantic Web Knowledge Base Artificial Intelligence Expert System

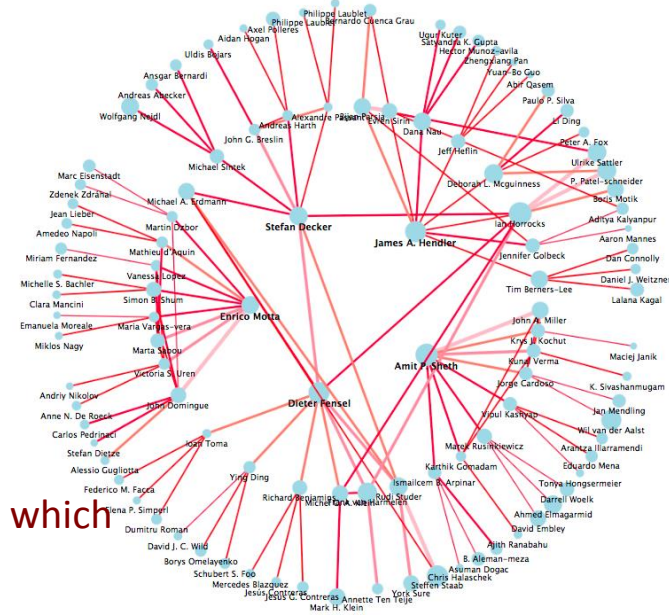
Zoom 5y 10y 15y 20y 30y All





# Rexplore

- Provides a variety of ways to explore scholarly data, e.g.:
  - To identify interesting research trends
    - E.g., which areas are expanding, which ones are shrinking, which interesting 'migration patterns' exist between areas, etc.
  - To discover interesting connections and similarities between authors beyond simple co-authorship (e.g., authors who share a similar academic trajectory)
  - To support fine-grained academic expert search
- Exploits a novel algorithm, Klink, which integrates semantics and statistics to automatically mine semantic relations between research topics from very large corpora and then uses these to support knowledge-based exploration



- World Wide Web +
- cT - Semantic Web -
- bg - Semantic Web Technology -
- bg - Semantic Web Rule Language
- bg - Web of Data
- cT - Semantic Technologies +
- cT - Semantic Search
- cT - Semantic Metadata
- cT - Social Web
- cT - Linked Open Data
- bg - Semantic Web Service +
- bg - Semantic Annotation -
- bg - Semantic Metadata
- bg - Semantic Wiki

# EventMedia Live – 2012 SW Challenge Winner



All Events    Featured Events

**Victor Wooten Band & Jimmy Herring Band**  
@ Wilbur Theatre, in 7 hours  
1 attendees

**Utah Jazz at Boston Celtics**  
@ TD Garden, in 7 hours

**Victor Wooten Band & Jimmy Herring Band**  
@ Wilbur Theatre, in 7 hours

**Jimmy Herring - VWB**  
@ Wilbur Theatre, in 7 hours  
Spring Tour

## David Karger's post about ISWC 2013

*The more seriously funny experience was at the ISWC demo session. **The two demos that most impressed me were systems for (i) browsing upcoming events (concerts etc.) and (ii) browsing academics and their publications.** Both of these systems were characterized by rich data models and nicely designed user interfaces that delivered valuable information and insights from their chosen domains.*



## David Karger's post about ISWC 2013



*The more seriously funny experience was at the ISWC demo session. **The two demos that most impressed me were systems for (i) browsing upcoming events (concerts etc.) and (ii) browsing academics and their publications.** Both of these systems were characterized by rich data models and nicely designed user interfaces that delivered valuable information and insights from their chosen domains.*

***The funny part is that neither of these applications should really be called a “Semantic Web application.”** Someone unaware of the Semantic Web ...would see a traditional data management and visualization problem that they would solve using traditional database tools (SQL) and web APIs. The fact that these tools are storing their data in a triple store instead of a SQL database is irrelevant to the user experience. And the fact that at least one of them is exposing a SPARQL endpoint for querying the data they are managing is good citizenship, helpful to the next project, but not important for this one.*

## There is more to a SW app than triple stores...

Now, funnily enough I do not necessarily agree with David that Rexplore (or for that matter, EventMedia Live) is not a SW application, however I basically agree with the essence of its comment:

***The fact that these tools are storing their data in a triple store instead of a SQL database is irrelevant to the user experience. And the fact that at least one of them is exposing a SPARQL endpoint for querying the data they are managing is good citizenship, helpful to the next project, but not important for this one.***

One would certainly think there is more to a SW application than simply using a triple store!

A new application award

# Semantic Web Challenge

<http://challenge.semanticweb.org>

[Home](#) | [Challenge 2012](#) | [Criteria](#) | [Submissions](#) | [Finalists](#) | [Winners](#) | [Former Challenges](#) | [Sponsors/Contact](#)

## Semantic Web Challenge 2012 - Call for Participation

Submissions are now invited for the 10th Semantic Web Challenge, the premier event for demonstrating practical progress towards achieving the vision of the Semantic Web. The 10th Semantic Web Challenge which will take place at the [11th International Semantic Web Conference](#) in Boston, USA.

As in previous years, the Semantic Web Challenge will consist of two tracks: the **Open Track** and the **Billion Triples Track**. The key difference between the two tracks is that the Billion Triples Track requires the participants to make use of the data set that has been crawled from the Web and is provided by the organizers. The Open Track has no such restrictions. The Challenge is open to everyone from industry and academia. The authors of the best applications will be awarded prizes and featured prominently at special sessions during the conference.

### Important Dates

- **Friday October 12, 2012, 23:59 CET:** Submissions due
- November 13-15: Semantic Web Challenge takes place at ISWC 2012

### Challenge Criteria

The Challenge is defined in terms of minimum requirements and additional desirable features that submissions should exhibit. The criteria for the Semantic Web Challenge 2012 are described on the [Criteria](#) page.

# Discussing Rexplore with Jim Hendler



# Discussing Rexplore with Jim Hendler

*Enrico, how come  
you stopped doing  
SW research?*





# Discussing Rexplore with Jim Hendler

*Jim, are you kidding?  
Rexplore follows exactly  
your philosophy: A little  
semantics goes a long way*



# Discussing Rexplore with Jim Hendler

*Yes, and **no semantics** goes even further*



## Bottom Line

- There certainly seems to be a clash of intuitions about what SW applications are about and, given the relatively advanced level of maturity of the field, it may be a good idea to try and clarify:
  - The nature of SW applications
  - The value proposition offered by the use of semantics on the web

# Disclaimer/Caveat #1

- I am not a philosopher, so I am not interested in discussing “What does it mean to be semantic” in an abstract way.
- My aim is to give pragmatic, engineering-oriented answers to these questions

*There is only one sound philosophical position on semantics*

*Is that “A little semantics goes a long way”?*



## Disclaimer/Caveat #2

- I am not going to attempt to ‘define’ the discipline!

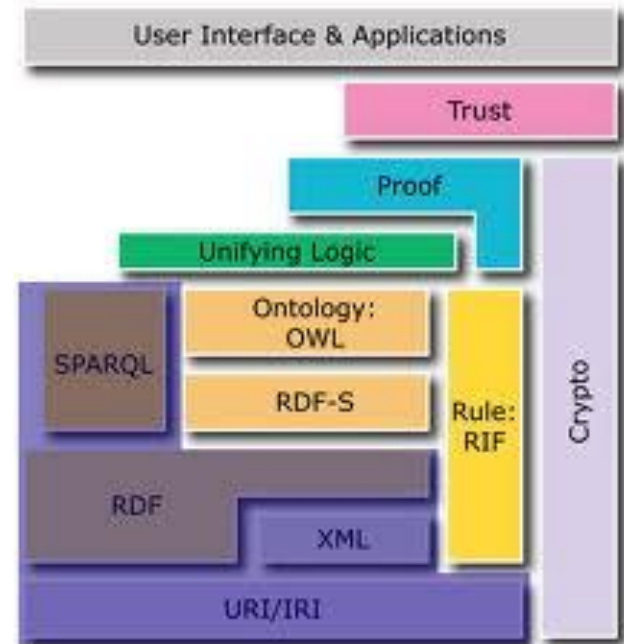
***“Disciplines are recognized, not defined”***

*Herb Simon*

- So, it is less a matter of defining SW research than highlighting the key principles of what we are doing to ensure that we capitalize on our strengths and build systems that push the envelope

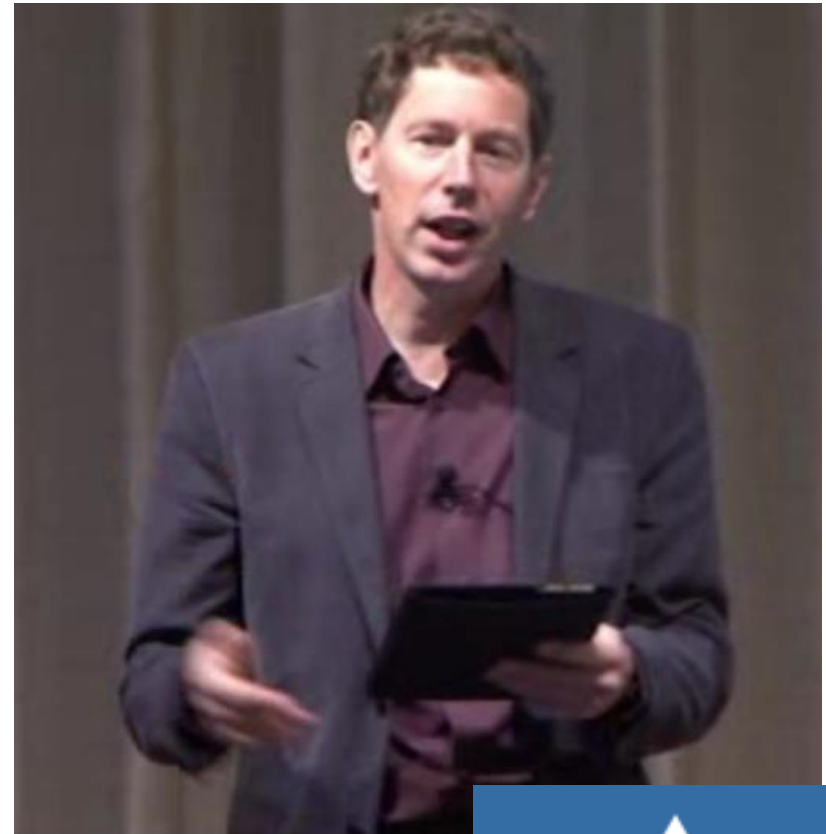
## Disclaimer/Caveat #3

- I will present one particular viewpoint drawing on my background in knowledge-based systems
  - This is just one perspective, others are also possible
- In particular, my **Knowledge-level** (Newell, 1982) perspective focuses on the role of semantics as an enabler of intelligent functionalities
  - Abstracts from specific symbol-level architectures and data structures
  - Takes for granted that there is lots of excellent work on specific infrastructure issues, e.g., ontologies, ontology and data integration, Web KR, query languages, large-scale reasoning, etc....



## Frank's talk at ISWC 2011

- Aims to identify “laws about the information universe”
- Several principles emerge
  - Factual knowledge is a graph
  - Terminology is hierarchical and exhibits low complexity
  - Small Hierarchy/Big Graph
  - Heterogeneity is unavoidable but solvable
- Focus primarily on the data side (SW as massive heterogeneous information space) in contrast with our focus on applications and the role of semantics



# Knowledge-Based System Perspective

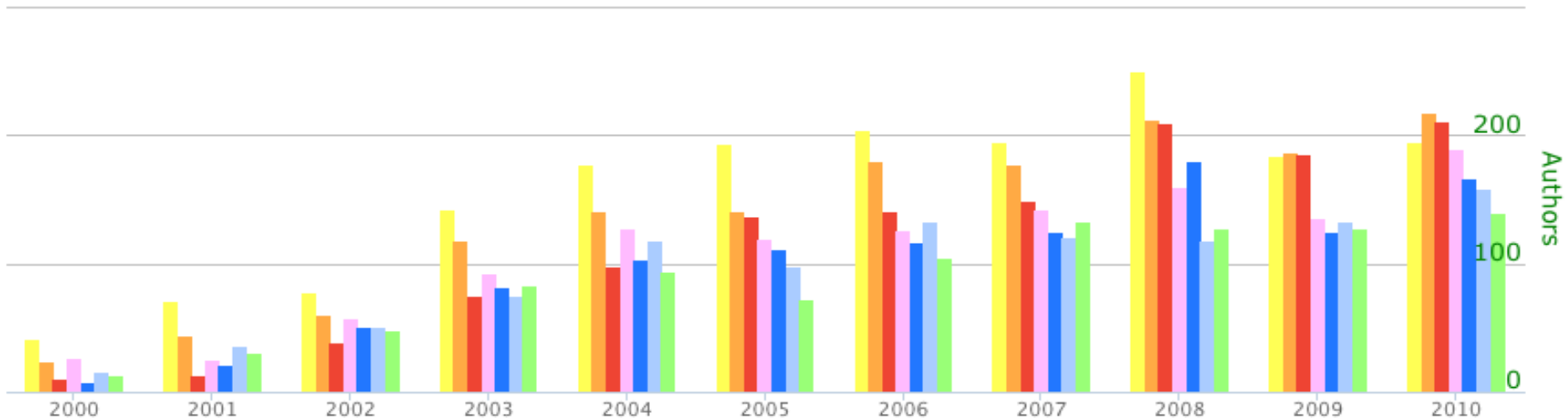
- My background is knowledge engineering and indeed one of the really cool aspects that attracted me to the SW was that it seemed to me it provided an opportunity to address fundamental AI issues in an exciting new, open, and distributed context, thus moving away from classic, 'closed' KBS

## Semantic Web

Main author inflow

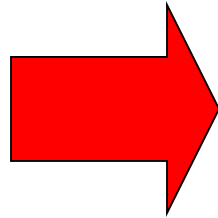
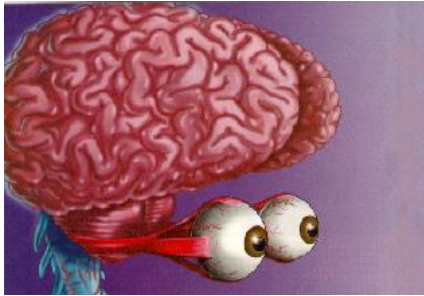


Zoom

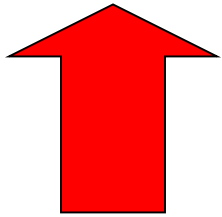




# Knowledge Representation Hypothesis in AI



***Intelligent Problem Solving***

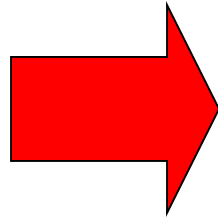
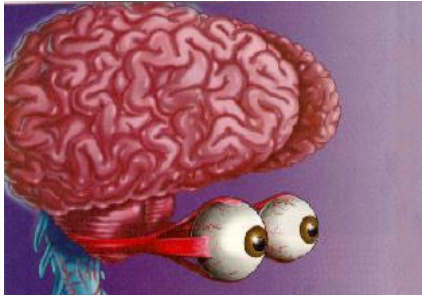


***Body of  
Knowledge***

Emphasis on causal and essential relationship between an explicit body of knowledge and the resulting intelligent behaviour

*Brian Smith, 1982*

# Task-centric view of Symbolic Intelligent Systems



## *Intelligent Problem Solving*

*Classification*

*Diagnosis*

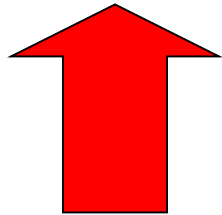
*Planning*

*Design*

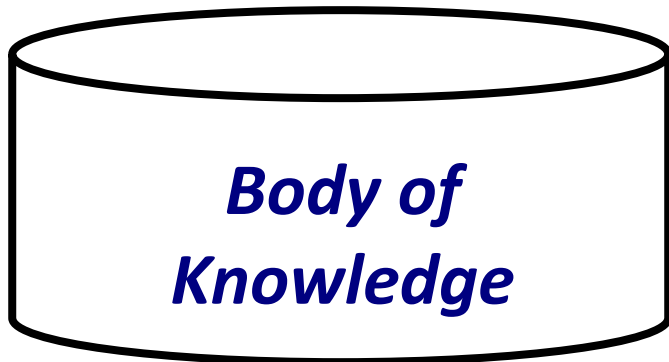
*Scheduling*

*Monitoring*

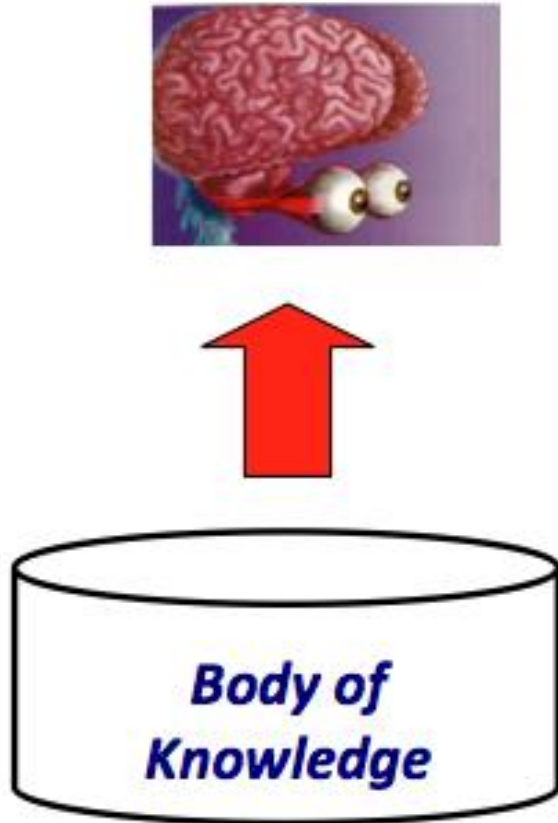
*Etc...*



*Body of  
Knowledge*



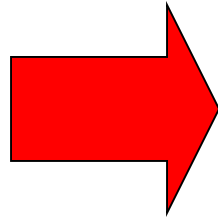
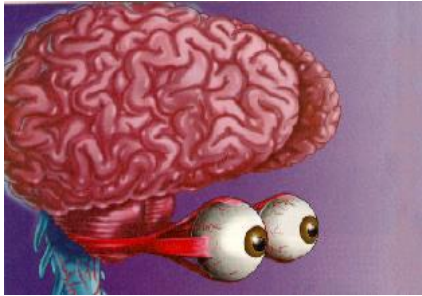
# Criteria for recognising intelligent knowledge-based systems



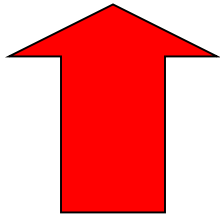
*Intelligent Problem Solving*

- Explicit representation of knowledge
- Causal link between body of knowledge and problem solving behaviour
- Task-specific criterion
  - Systems tackling tasks complex enough to require knowledge-based problem solving

# Intelligent Semantic Web Systems (?)



***Intelligent Problem Solving***

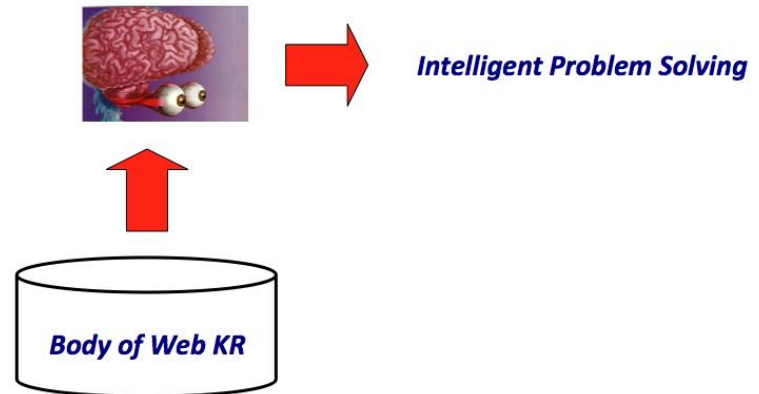


***Body of Web KR***



# Intelligent Semantic Web Systems (?)

- Several problems with this simple view
  - Semantics on the web is not the same as web semantics
    - Putting a classic, closed KBS on the web does not add anything new
  - Also, in contrast with the AI scenario, it is difficult to come up with a catalogue of tasks defining the range of intelligent activities that SW systems can support in the same way as it was done in the KBS community.
  - SW tasks appear to be more adhoc in nature
    - Analysis by Van Harmelen et al (KCAP 2009) identified the following tasks:
      - Search
      - Browsing/Querying
      - Data Integration
      - Personalization
      - Service Selection
      - Service Composition
      - Semantic Enrichment



## Key postulates

*Just having a body of Web-KR up on the web to drive some functionality in a manner analogous to traditional, closed KBS is not enough to talk about a 'real' SW application*

*The intuition is that SW applications ought to have at least elements to do with openness to other data sources and/or scale (+ all the other elements that derive from these two)*

# Toward a New Generation of Semantic Web Applications

Mathieu d'Aquin, Enrico Motta, Marta Sabou, Sofia Angeletou, Laurian Gridinoc, Vanessa Lopez, and Davide Guidi, *Open University*

*IEEE Intelligent Systems* 23(3),

pp. 20-28,

May/June 2008

*A new generation of applications offers insight into the Semantic Web's current and future challenges—as well as the opportunities it might provide for users and developers alike.*

**A**lthough research on integrating semantics with the Web started almost as soon as the Web was in place, a concrete Semantic Web—that is, a large-scale collection of distributed semantic metadata—emerged only over the past four to five years.

The Semantic Web's embryonic nature is reflected in its existing applications. Most of

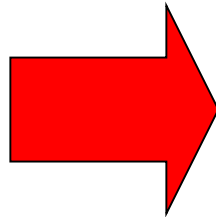
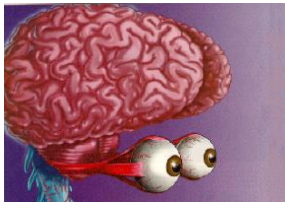
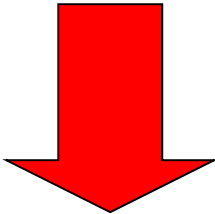
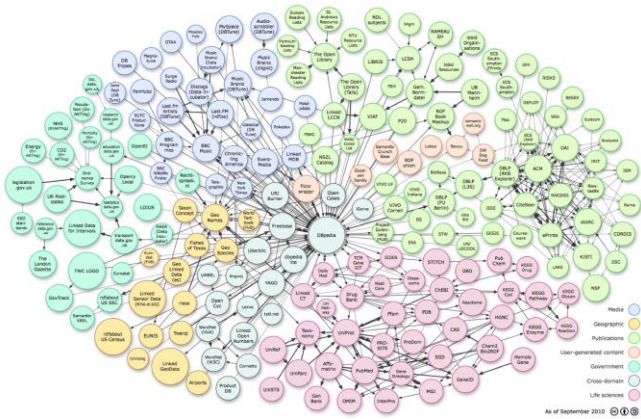
these applications tend to produce and consume their own data, much like traditional knowledge-based applications, rather than actually exploiting the Semantic Web as a large-scale information source.<sup>1</sup>

These first-generation Semantic Web applications<sup>1</sup> typically use a single ontology that supports integration of resources selected at design time. An early influential example from the academic world is CS Active Space (<http://cs.aktivespace.org>). This application combines data about UK computer science research from multiple, heterogeneous sources (such as databases, Web pages, and RDF data) and lets users explore the data through an inter-

cessing. Corporate Semantic Web application areas include the car industry (such as Renault's system for managing project history), the aeronautical industry (such as Boeing's use of semantic technologies to gather corporate information), and the telecommunication industry (such as British Telecom's system for enhancing digital libraries).

Although corporate Semantic Webs often provide perfectly adequate solutions to a company's needs, they actually fall short of fully exploiting the Semantic Web's exciting potential as a large-scale source of background knowledge. To address this, we began an ambitious research program two years ago dubbed "Next-Generation Semantic

# Knowledge Engineering in the age of the SW



Systems which exploit large-scale, distributed knowledge to do interesting stuff

***Intelligent Behaviour***



# Features of NGSW Systems (vs Classic KBS)

	<b>Classic KBS</b>	<b>NGSW Systems</b>
<b>Ontology/Data Selection</b>	Static/Design-time/Closed	Dynamic/Run-time/Open
<b>Provenance</b>	Centralized	Distributed
<b>Size</b>	Small/Medium	Very Large
<b>World View</b>	Homogeneous	Heterogeneous
<b>Repr. Schema</b>	Homogeneous	Heterogeneous
<b>Quality of data</b>	High	Very Variable
<b>Degree of trust</b>	High	Very Variable

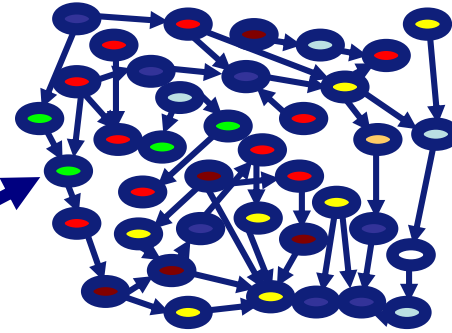
# Major shift in the type of reasoning

	<b>Classic KBS</b>	<b>NGSW Systems</b>
<b>Intelligence</b>	A function of applying sophisticated logical reasoning to tackle complex real-world tasks	A side-effect of the ability to reason with large scale data and heterogeneous quality and representation, typically integrating (a bit of) logical reasoning with other technologies –e.g., statistics and NLP

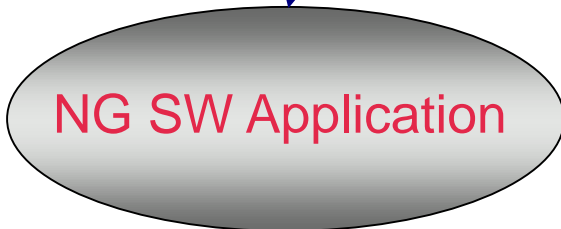


## **NGSW Applications: Realising the research programme**

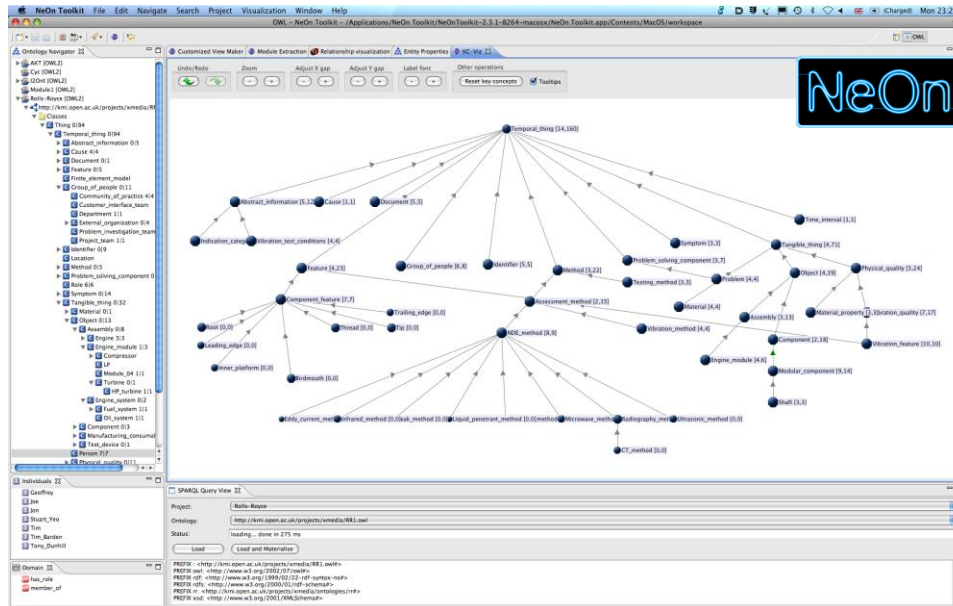
# Watson: A Gateway to the Semantic Web



- Semantic Web Search Engine
- Development Platform for NGSW Applications
- Research Platform for studying the Semantic Web



# Watson-based Applications

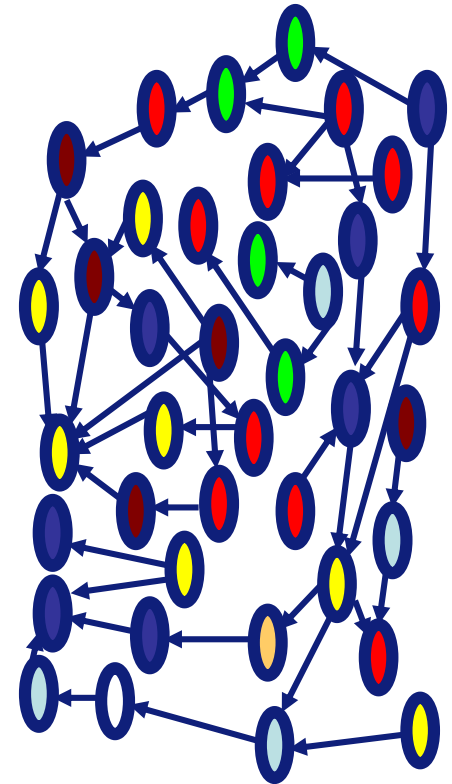
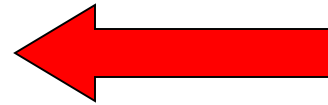


# Folksonomy Tagspace Enrichment



## All time most popular tags

06 africa amsterdam animals architecture art august australia autumn baby barcelona  
**beach** berlin **birthday** black blackandwhite blue boston bw **california**  
 cameraphone camping canada canon car cat cats chicago china christmas  
 church city clouds color concert day dc dog england europe fall **family**  
 festival film florida **flower flowers** food france **friends** fun garden  
 geotagged germany girl graffiti green halloween hawaii hiking holiday home  
 honeymoon hongkong house india ireland island **italy japan** july june kids lake landscape  
 light live **london** losangeles macro may **me** mexico mountain mountains museum music  
**nature new newyork** newyorkcity newzealand night nikon nyc ocean october **paris**  
 park **party** people portrait red river roadtrip rock rome san **sanfrancisco** school  
 scotland sea seattle september show sky snow spain spring street **summer** sun  
 sunset sydney taiwan texas thailand tokyo toronto **travel** tree trees trip uk urban  
**usa vacation** vancouver washington water **wedding** white winter yellow  
 york ZOO

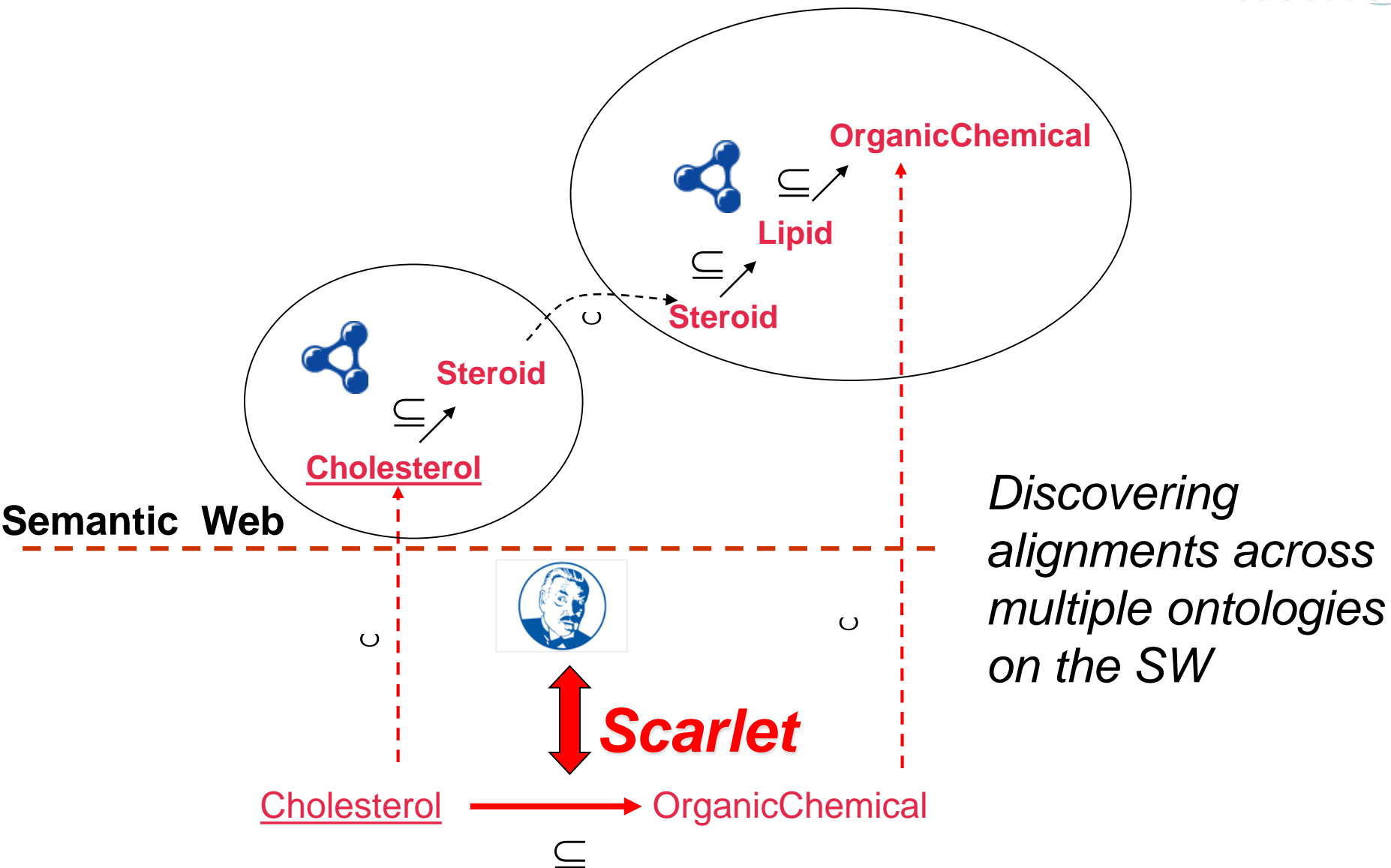




# Using the SW as a source of background knowledge for Ontology Alignment

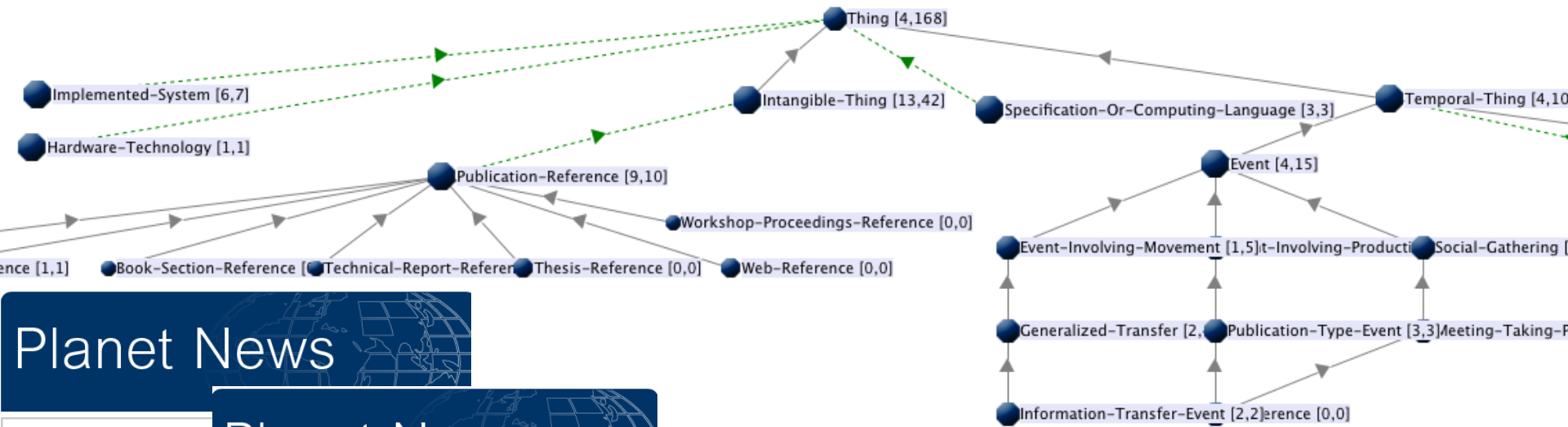


# Strategy





# Ontology Evolution using Background Knowledge



**Planet News**

**KMi scoops ASWC Best Paper award for second year running**  
 KMi Reporter, Monday 21 December 2010

**SOA4All: Leading the Future in Services**  
 Carlos Pedrinaci, Monday 10 May 2010

**SAP Says "Future Internet: Right Here, Right Now"**  
 John Domingue, Thursday 29 April 2010

**SOA4All Logo**

**SOA4All**

Last week saw SAP Barcelona. Drive reviews, this time exploitation of pre expectations the project.

Yesterday saw the end of the highly stimulating day and a half International Research Forum (IRF) hosted by SAP at the Lindner Hotel & Sporting Club Wiesensee. The annual invitation-only forum brings together the leading thinkers from academia and industry around a preselected topic and the theme of this year's event was "Future Internet: Right Here, Right Now".

The event was led by Lutz Heuser SAP Vice President for Research and was structured around visionary and challenge talks, panels and break out sessions. A significant SAP support team for the event incorporated several reporters, a press photographer, a film crew, live concept/mind mappers, sound recordists and professional moderators. Besides Lutz other notable attendees included: Claudia Alsdorf CEO of Original1 (a new cloud computing company created by SAP, Nokia and Giesecke & Devrient), Mario Campolongo Director of the Emerging Technologies and Infrastructures, Director of DG-INFO, Martin Curley Director of Intel Labs Europe, Claudia Funke a Senior Partner at McKinsey & Company, Matthias Kaiserswerth Head of IBM Research Strategy in Systems Management and Compliance, Peter Kukrlick Chief Product Officer and Executive Board Member Software AG, Latif Ladid President of the Internet Protocol Version 6 (IPv6) Forum, Maximilian Metzger Director of Information and Communication Technologies, New Services at the German Federal Ministry of Education and Research, Joachim Schaper Vice President at SAP Research and Wolfgang Wahster head of DFKI (and also now head on the winning European Institute of Innovation and Technology Knowledge Innovation Community).

Thus, with renewed strength and motivation, the of research and development, in which it is expected the evolution of services on the Web.

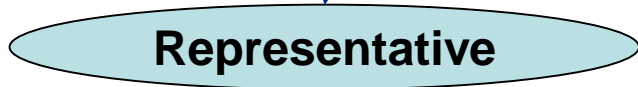


# Relation Discovery Examples

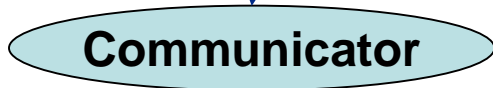
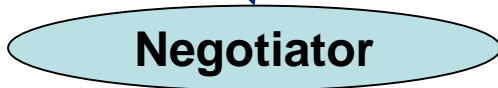
*New  
 Concepts*



*WordNet*



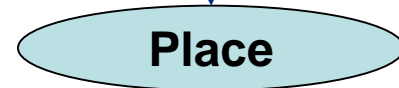
*Background  
 Knowledge*



*Ontology  
 Concepts*



*Scarlet*



$\sqsubseteq$

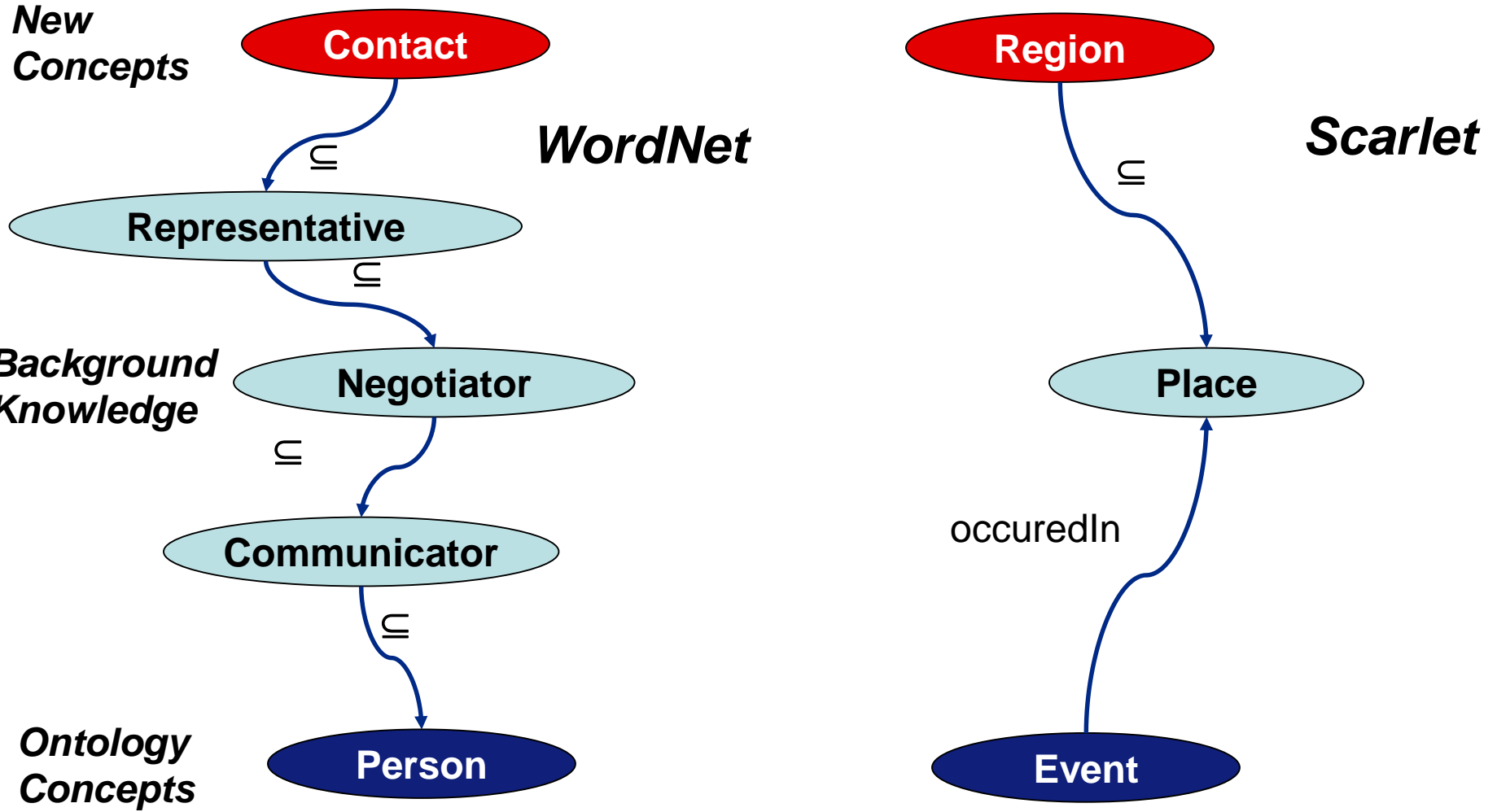
$\sqsubseteq$

$\sqsubseteq$

$\sqsubseteq$

$\sqsubseteq$

occuredIn



# NGSW Applications

- Systems able to exploit distributed semantic information at very large scale, (typically) by combining semantic and statistical methods and tackling the challenges associated with operating in a massively distributed and heterogeneous environment
- **Focus on openness and scale with respect to using large amounts of massively distributed SW sources**
- Essentially reflect the view of the SW as “Knowledge Medium”

## **The Next Knowledge Medium**     *AI Magazine 7(1), 1986*

Mark Stefik

*Intelligent Systems Laboratory, Xerox Palo Alto Research Center, 3333 Coyote Hill Road,  
 Palo Alto, California 94304*

*“An information network with semi-automated services for the generation, distribution, and consumption of knowledge”*

# Rexplore

Making Sense of Scholarly Data

Info Migrations

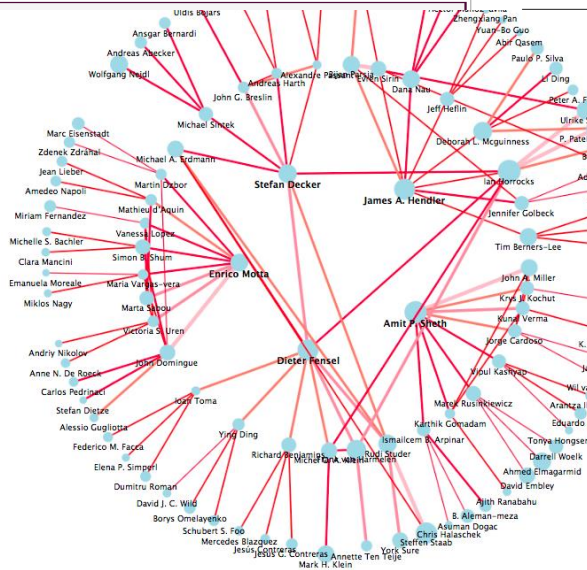
## Semantic Web

Publications: 22 143  
Citations: 120 704

- Plot authors and publications
- Plot average citations vs authors and publications
- Explore authors
- Explore publications

## Semantic Relationships

- Reload Semantic Web relationships
- Deactivate semantic layer

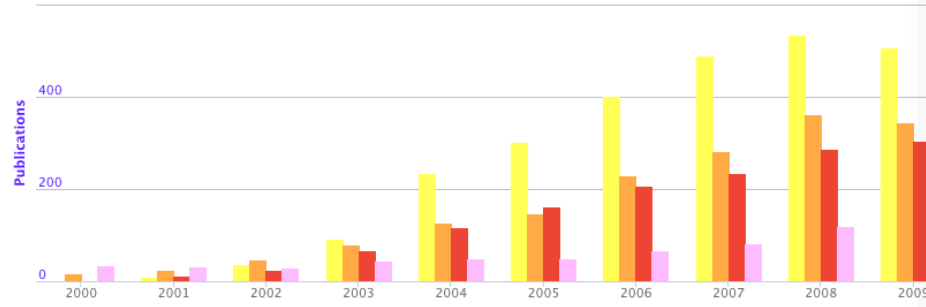


## Semantic Web

broaderGeneric Topics - publication trend

Semantic Web Service Semantic Annotation Semantic Web Technology Linked Data

Zoom 5y 10y 15y 20y 30y All



World Wide Web

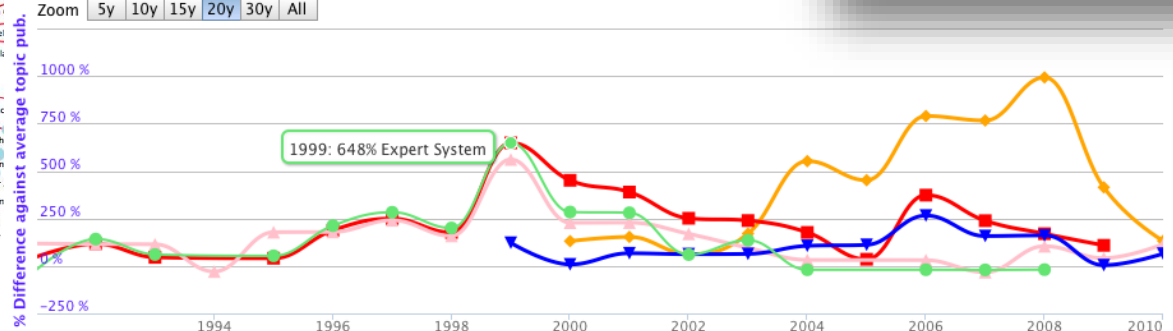
- cT - Semantic Web -
- bG - Semantic Web Technology -
- bG - Semantic Web Rule Language
- bG - Web of Data
- cT - Semantic Technologies +
- cT - Semantic Search
- cT - Semantic Metadata
- cT - Social Web
- cT - Linked Open Data
- bG - Semantic Web Service +
- bG - Semantic Annotation -
- bG - Semantic Metadata
- bG - Semantic Wiki

## Enrico Motta

Publications normalized by topics

total publications Semantic Web Knowledge Base Artificial Intelligence Expert System

Zoom 5y 10y 15y 20y 30y All



# Motivations for Rexplore

- Some limitations of current tools
  - Poor support for ‘horizon’ tasks, concerned with making sense of what goes on in research, as opposed to searching for specific authors and publications
  - Poor support for fine-grained academic search –e.g., parametrized with respect to fine-grained expertise, degree of experience, coverage of multiple research areas, geography, academic ranking criteria, etc
  - Poor support for discovering relations and similarities between researchers, beyond obvious co-authorship relations
  - Poor HCI, loss of context during navigation, etc.
  - Etc...

# Limitations of existing tools (lack of semantics!)

- In order to find experts in an area, e.g., Semantic Web, I need systems which have some level of understanding that the Semantic Web is a research area
- But understanding what is a research area also means understanding what is not a research area
  - E.g., “case study” is often used as a tag for papers, but it is not actually a research area
- In addition, research areas have a structure.
  - For instance, knowing which are the sub-areas of the SW would allow a system to provide more precise answers about the competence of the relevant researchers and the important trends within the SW
  - It is also the case that multiple labels can be used for the same research area
    - E.g., “Ontology Alignment” and “Ontology Matching” are labels for the same research area

Co-authors (217)

- John Domingue
- Victoria S. Uren
- Marta Sabou
- Mathieu d'Aquin
- Vanessa Lopez



Conferences (45)

- ISWC
- EKAW
- ESWS
- K-CAP
- ASWC

Journals (39)

- IJMMS
- Journal of Web Semantics
- EXPERT
- PROC INST CIVIL ENG-GEOTECH E
- KER

Keywords (433)

- Case Study
- Information Extraction
- Knowledge Acquisition
- Knowledge Based
- System Knowledge
- Engineering Knowledge
- Management
- Large Scale
- Natural Language
- Ontology

Academic > Authors > Enrico Motta

Embed Subscribe



**Enrico Motta**

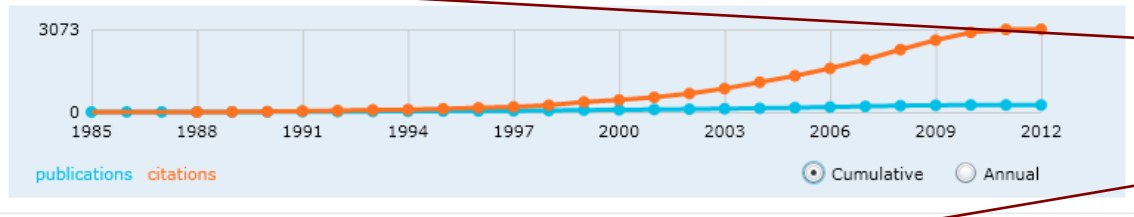
Open University UK

Publications: 324 | Citations: 4193

Fields: Artificial Intelligence, World Wide Web, Databases

Collaborated with 217 co-authors from 1985 to 2011 | Cited by 4550 authors

Edit



Publications (324) Export

Sort by: Year

[Watson, more than a Semantic Web search engine](#) (Citations: 2)

Mathieu d'Aquin, **Enrico Motta**

Published in 2011.

[Ontology augmentation combining semantic web and text resources](#)

Miriam Fernandez, Ziqi Zhang, Vanessa Lopez, Victoria Uren, **Enrico Motta**

Conference: International Conference on Knowledge Capture - K-CAP, pp. 9-16, 2011

[Extracting relevant questions to an RDF dataset using formal concept analysis](#)

Mathieu d'Aquin, **Enrico Motta**

Published in 2011.

[Data Linking: Capturing and Utilising Implicit Schema-level Relations](#) (Citations: 3) View...

Andriy Nikolov, **Enrico Motta**

Published in 2010.

[Reflections on five years of evaluating semantic search systems](#) (Citations: 3)

Victoria S. Uren, Marta Sabou, **Enrico Motta**, Miriam Fernández, Vanessa Lopez, Yuanguai Lei

Journal: International Journal of Metadata, Semantics and Ontologies - IJMSO, vol. 5, no. 2, pp. 87-98, 2010

Very high level research fields

Only co-autorship is provided

Old name for IJHCS (changed long ago!)

This journal has nothing to do with my research areas

Case Study is not a research area

KB and KBS are the same research area

# Semantic Analysis

- No explicit notion of Research Area, aside from very abstract topics, such as Computer Science, WWW, etc.
- No semantic relations between research areas
- Homogenous treatment of keywords, regardless of whether they denote research areas or other entities (e.g., project, type of study, application area, etc.)
- Only a limited set of static relations between authors
- Limited degree of disambiguation between homonymous authors
- Many co-reference errors concerning authors
  - i.e., situations where distinct authors are actually the same person
- No disambiguation between homonymous areas
  - E.g., linked data!



# Identifying and linking research topics with **Klink**

- The Klink algorithm automatically identifies research topics in the literature, mines semantic relations between research topics and uses the resulting data to support knowledge-based exploration, pattern extraction and author clustering in Rexplore.
- Three types of semantic relations are currently mined:
  - *Skos:broaderGeneric* (A, B) – A is a sub-area of B.
    - E.g., “Semantic Web Services” is a sub-area of “Web Services”
  - *relatedEquivalent* (A, B) – A and B are normally used to denote the same research area. This is defined as a sub-property of *skos:related*.
    - E.g., “Ontology Matching” and “Ontology Mapping” denote the same area
  - *contributesTo* (A, B) – The outputs from area A are relevant to research in area B. In other words, if I am interested in area B, I may also examine papers/authors in area A. This is also defined as a sub-property of *skos:related*.
    - E.g., Research in “Ontology Engineering” contributes to research in “Semantic Web”

## Semantic Relationships ■

### ⊗ Reload Ontology Integration relationships

- ⊗ Problem Solving +
  - ⊗ bG - Domain Knowledge +
    - ⊗ cT - Ontology +
      - ⊗ bG - Ontology Mapping -
        - ⊗ bG - **Ontology Integration**
      - ⊗ bG - {Ontology Matching, Ontology Alignment} -
        - ⊗ cT - **Ontology Integration**
  - ⊗ cT - Knowledge Base +
    - ⊗ bG - Knowledge Representation +
      - ⊗ cT - Knowledge Acquisition +
        - ⊗ bG - Domain Knowledge +
    - ⊗ cT - Expert System +
      - ⊗ cT - Knowledge Acquisition +
  - ⊗ cT - Artificial Intelligence +
    - ⊗ bG - Knowledge Representation +
    - ⊗ cT - Domain Knowledge +
    - ⊗ cT - Expert System +
- ⊗ World Wide Web +
  - ⊗ cT - Semantic Web +
    - ⊗ bG - Semantic Interoperability +
      - ⊗ cT - Ontology Mapping +
      - ⊗ cT - {Ontology Matching, Ontology Alignment} +
    - ⊗ cT - Ontology +
      - ⊗ bG - Ontology Mapping +
      - ⊗ bG - {Ontology Matching, Ontology Alignment} +
- ⊗ Natural Language +
  - ⊗ cT - Knowledge Representation +

*About 1500 topics linked by  
almost 3000 semantic relationships*

# ACM and other similar classifications

## XII. Intelligent Web Services and Semantic Web

- I. Intelligent Web service languages
- II. Internet reasoning services
- III. Ontology design
- IV. Ontology languages

- These exhibit several major limitations:
  - The relations between entries are unclear
    - They are meant to be sub-areas, but for many of them it can be argued that they are not really sub-areas
  - The different types of relationships are not distinguished
  - Rather shallow
    - Most areas we know about are not listed – e.g., only 4 topics are classified under Semantic Web
  - Static, manually defined, hence they get obsolete very quickly

# Visualizing research trajectories

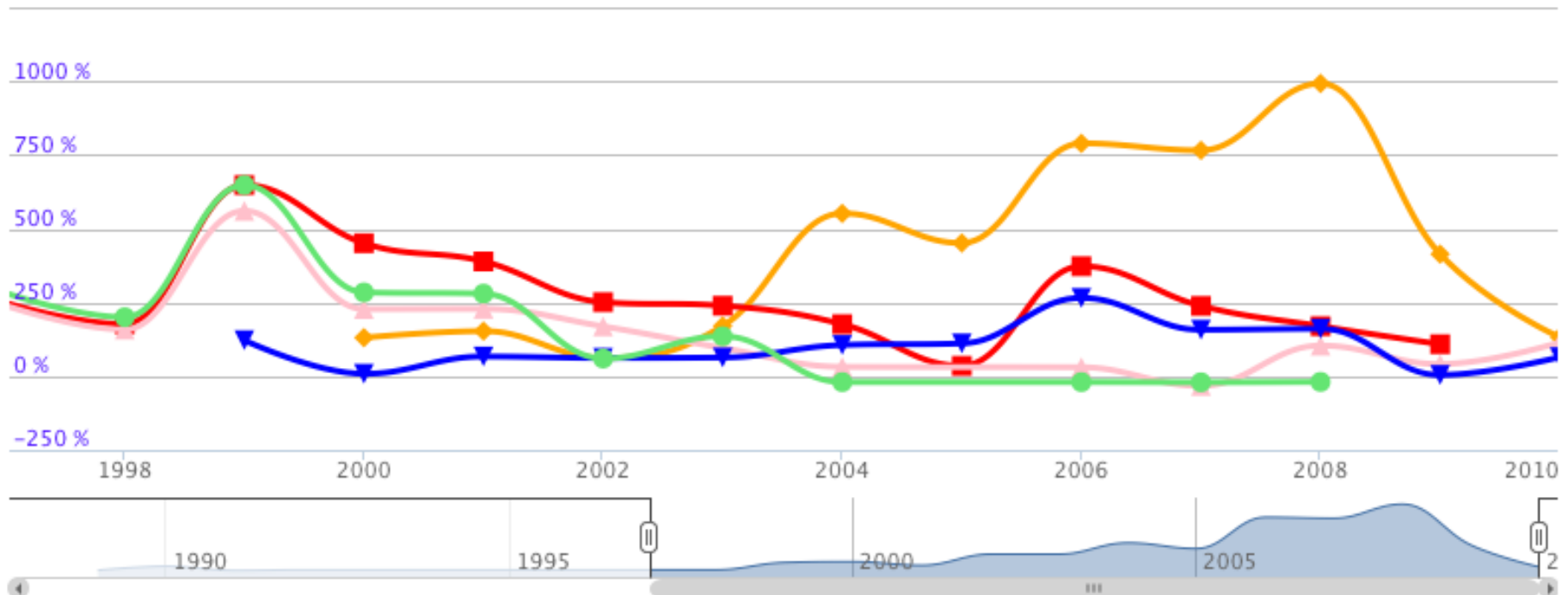
Enrico Motta  
Publications normalized by topics



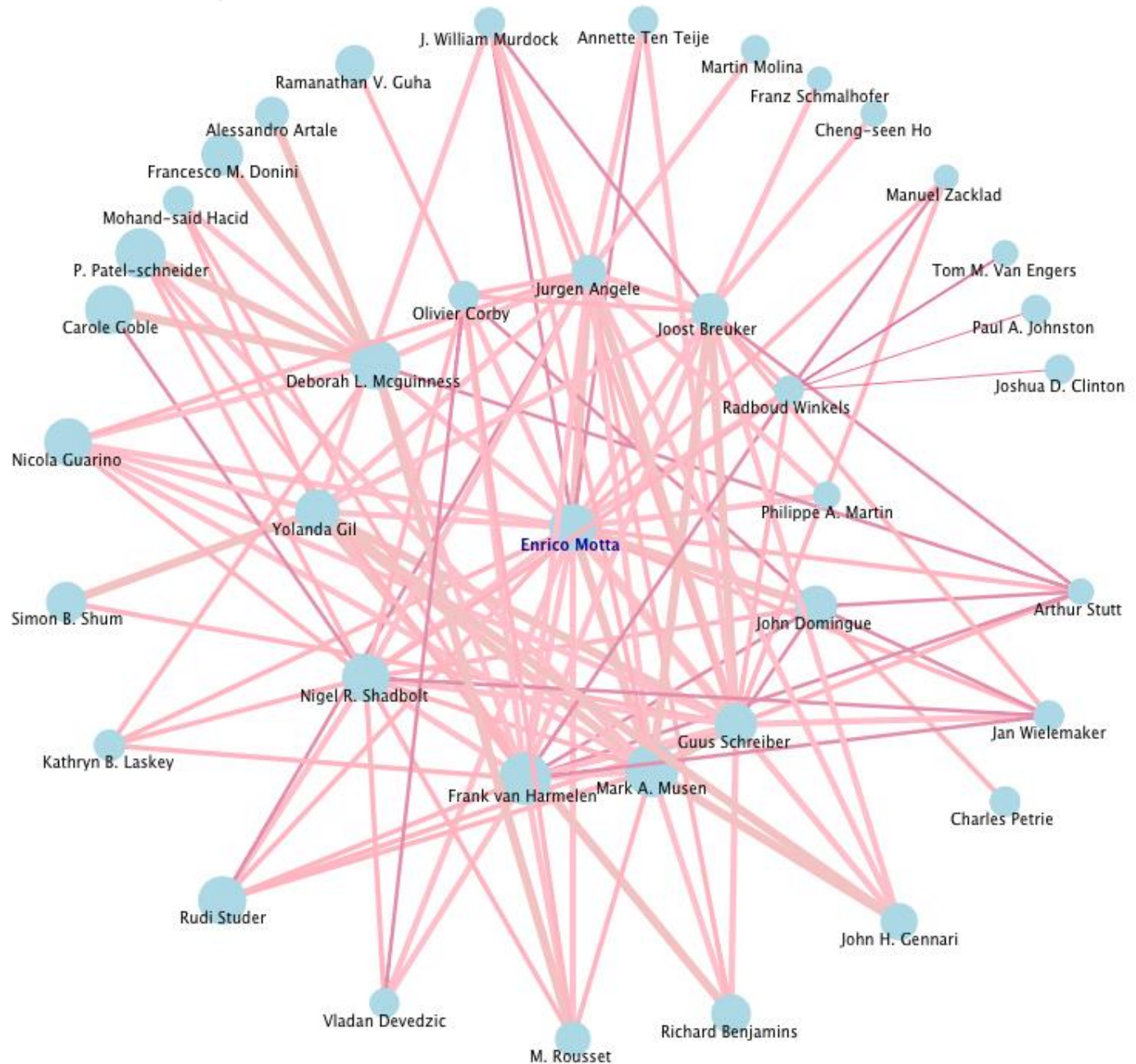
total publications Semantic Web Knowledge Base Artificial Intelligence  
Information Retrieval Expert System

Zoom 5y 10y 15y 20y 30y All

% Difference against average topic pub.



# Shared Research Trajectories



*The authors who are most similar to a particular author with respect to the evolution of their research interests over time.*

# Comparing publication trends

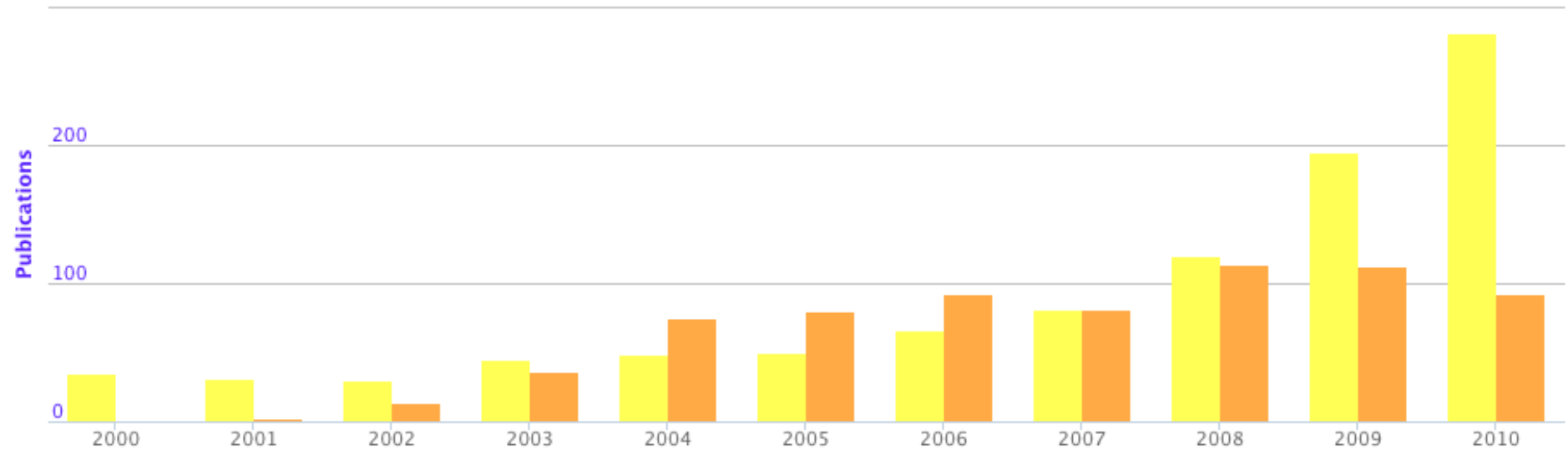
## Semantic Web

broaderGeneric Topics - publication trend



**Linked Data** **Web Ontology Language**

Zoom 5y 10y 15y 20y 30y All



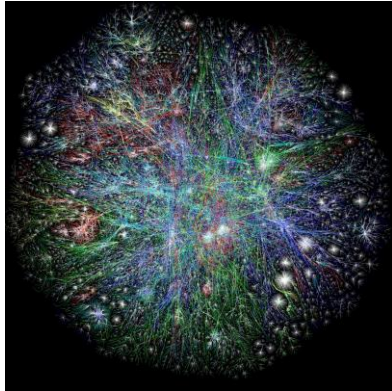
## Semantics as analytical engine

- Systems which like Rexplore are able to handle large amounts of distributed web data (typically) by combining semantic, statistical, and computational linguistics methods
- These systems use semantics as an analytical tool to identify distinctions and give meaning to regularities in the data to support problem-solving, sense-making or decision-making
- **Focus on openness and scale with respect to their ability to impose semantics (thus providing model-theoretical interpretations) on large amounts of massively distributed web data, which are not already formalised using SW standards**



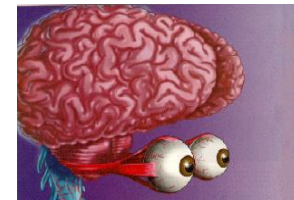
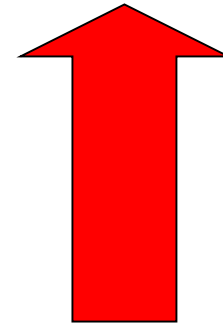
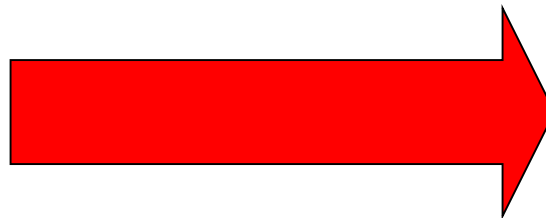
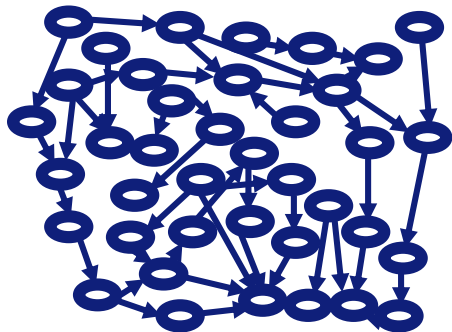
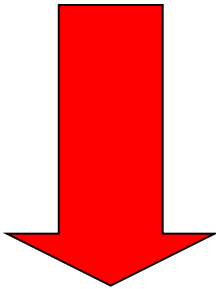


# A different class of SW applications (semantics as analytical engine)



Semantics as a medium to impose formal distinctions on large amounts of unstructured information to make explicit the underlying semantic patterns and enable intelligent functionalities

***Intelligent  
Functionalities***



# Carla Bruni

Singer-songwriter

Carla Bruni, now known by her name of use Carla Bruni-Sarkozy is an Italian-French singer, songwriter and former model. She married the then President of the French Republic Nicolas Sarkozy in February 2008.

Wikipedia



**Born:** December 23, 1967 (age 45), Turin, Italy

**Height:** 5' 9" (1.75 m)

**Spouse:** Nicolas Sarkozy (m. 2008)

**Children:** Giulia Sarkozy, Aurélien Enthoven

**Albums:** Quelqu'un m'a dit, No Promises, Comme si de rien n'était, Little French Songs, J'arrive à toi, Chez Keith et Anita

## Songs

<a href="#">Quelqu'un M'A Dit</a>	2003	<a href="#">Quelqu'un M'a Dit</a>
<a href="#">Chez Keith Et Anita</a>	2013	<a href="#">Chez Keith Et Anita</a>
<a href="#">L'amoureuse</a>	2008	<a href="#">Comme Si de Rien N'E...</a>
<a href="#">Raphaël</a>	2003	<a href="#">Quelqu'un M'a Dit</a>
<a href="#">Le Plus Beau du Quartier</a>	2003	<a href="#">Quelqu'un M'a Dit</a>

## People also search for



Nicolas Sarkozy  
Spouse



Valérie Trierweiler



Cécilia Sarkozy



François Hollande



Woody Allen

Non-Analytical

Analytical

*Implicit vs explicit semantic patterns*

# Conclusions

- Many of us share the intuition that SW applications ought to have elements to do with openness to data sources and scale –plus, optionally, others
- Our “NGSW Applications” research programme investigated the feasibility and the value derived from building applications exploiting large-scale distributed semantic structures
- However, semantics is also an effective tool to help users to make sense of large web data, acting as a device to introduce distinctions and give meaning to structures that may have been generated using other technologies (e.g., statistical methods)

KNOWLEDGE MEDIA

KMi

I N S T I T U T E

