



# Rexplore

## ***Making Sense of Research***

Enrico Motta

Knowledge Media Institute  
The Open University  
United Kingdom

# Hats I wear....

- Researcher
- Research Manager
- Supervisor/Mentor
- Editor-in-chief of a journal
- Advisor to strategic research programmes
- etc

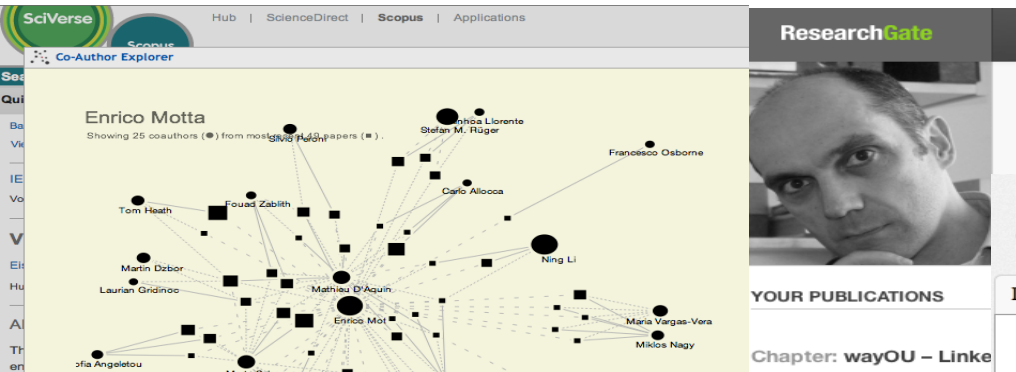


# Tasks



- **Academic Expert Search.**
  - E.g., “find me researchers with expertise in both Social Networks and Semantic Web, with at least some publications in CHI and ISWC, with more than 15 years research experience, a h-index greater than 15, etc”
- **Understanding Research Dynamics**
  - E.g., as EiC, I often need to make a decision about proposals for a special issue in a particular topic. This requires to understand whether the area is ‘hot’ right now or is decreasing in importance, who are the key people and groups, etc..

# Exploring scholarly data: a variety of options....



Academia.edu profile for Enrico Motta. The profile includes a search bar, a profile picture, and a list of research interests: Computer Science, Semantic Web, Human Computer Interaction, and 2 more. It also shows a list of papers, including "PlanetOnto: From News Publishing to Integrated Knowledge Management Support" and "Toward a New Generation of Semantic Web Applications".

AMiner search interface showing a search for "semantic web". The search results suggest "Semant Jain" as a related expert.

Do you mean: [Semant Jain](#)

Experts found for "semantic web". The list includes Enrico Motta (H-index: 37, #Papers: 181, #Citations: 5660) and Katia Sycara (H-index: 58, #Papers: 280, #Citations: 22669). Each entry includes a profile picture, name, and a list of associated keywords.

ResearchGate profile for Enrico Motta. The profile includes a profile picture, name, and a list of publications. The Mendeley logo is also visible.

My Library section showing a list of documents. The list includes "All Documents", "Recently Added", "Favorites", "Needs Review", "My Publications", "Unsorted", "AI", "Enterprise Modelling", "KR", "Large Scale Web Data", and "Machine Learning".

Enrico Motta profile on a platform. The profile includes a profile picture, name, and a list of titles: "Professor of Knowledge Technologies, KMI, The Open University", "Semantic Web - Ontology Engineering - Knowledge Systems", and "Verified email at open.ac.uk".

Citation indices and Citations to my articles. The citation indices table shows: Citations (8450), h-index (48), i10-index (158) for All, and Citations (5107), h-index (36), i10-index (119) for Since 2007. The bar chart shows citations to articles from 1990 to 2012, with a peak around 2007.

| Citation indices | All  | Since 2007 |
|------------------|------|------------|
| Citations        | 8450 | 5107       |
| h-index          | 48   | 36         |
| i10-index        | 158  | 119        |



Search results for "Enrico Motta" on a platform. The results show a list of publications with columns for Authors, Title, and Venue. The first result is "PowerAgent: Supporting users in querying and exploring the Semantic Web" by Vanessa Lopez, Miriam Fernandez, and Enrico Motta.

Google Scholar profile for Enrico Motta. The profile includes a search bar, a list of citations, and a list of co-authors. The list of co-authors includes John Domingue, Marta Sabou, and others.

# Lack of comprehensive and integrated support



“There is still a need for an *integrated solution*, where the different functionalities and visualizations are provided in a coherent manner, through an environment able to support a seamless navigation between the different views and functionalities”

*Dunne et al., 2012*

# Digital library perspective



- Tools tend to focus primarily on authors' publications and citations rather than sensemaking or expert search (in particular highly-faceted expert search)

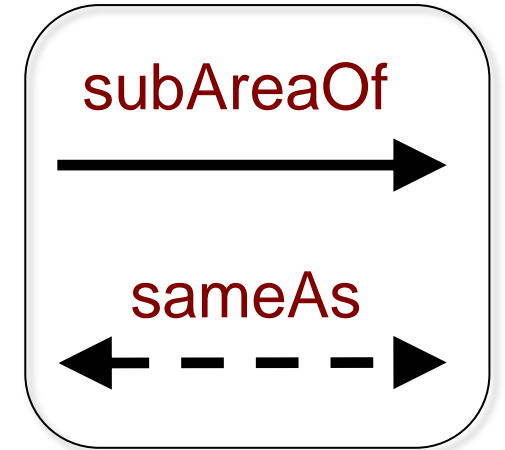
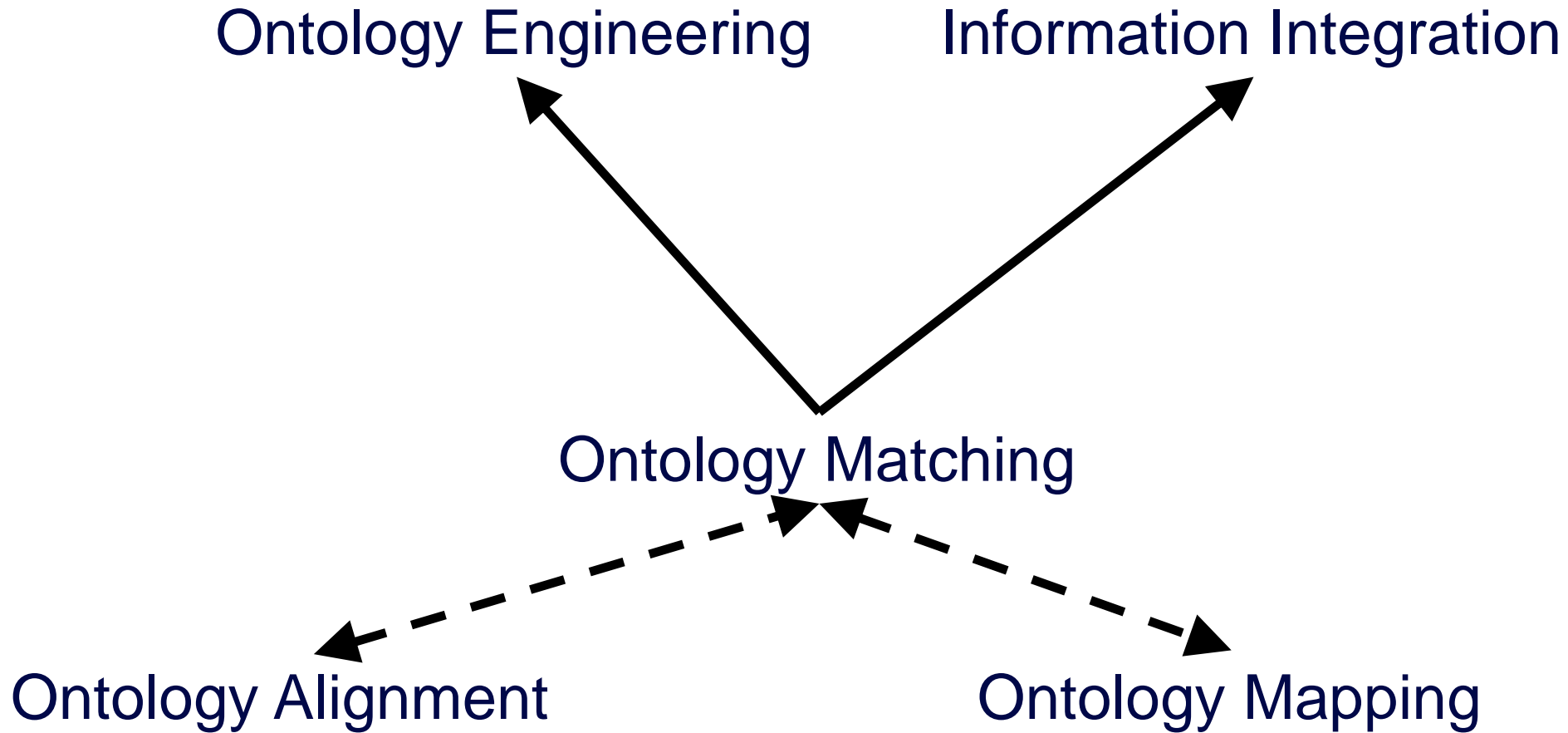


# Lack of a semantic treatment of research topics



- Current tools do not treat research topics as ‘first class citizens’.
  - E.g., a tool may support a keyword search for papers on Ontology Matching, but by and large tools does not ‘understand’ that Ontology Matching is actually a research area
- Crucially, 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

# Relations between research areas





# 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

- 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

# Rexplore

Exploring Scholarly Data



About Publications Team Contact

Info Migrations

## Semantic Web

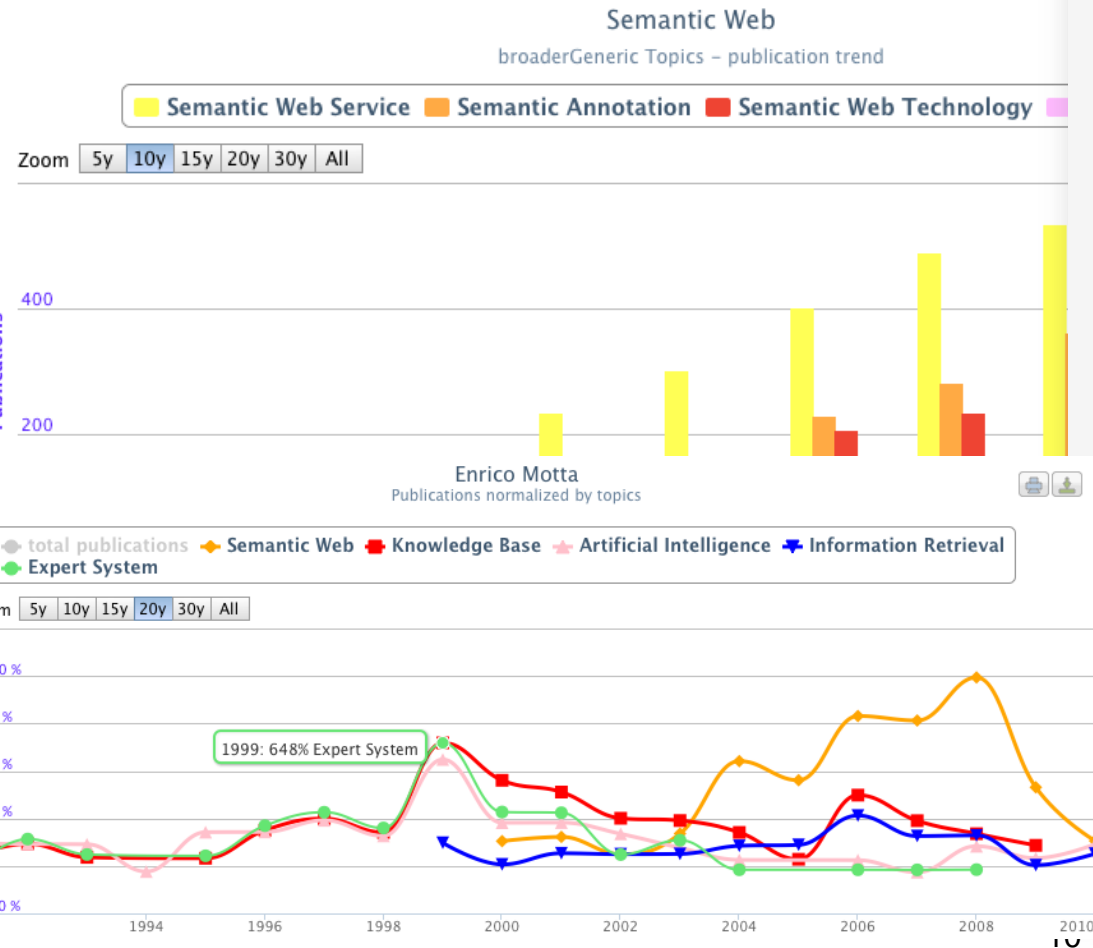
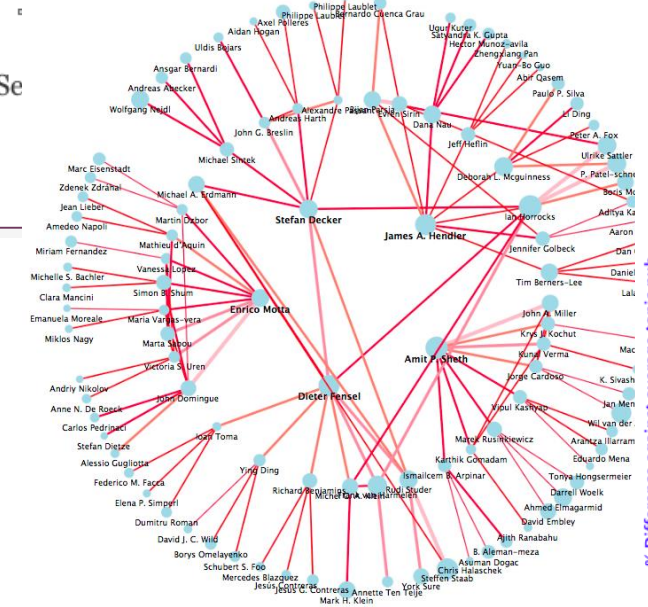
Publications: 22 143

Citations: 120 704

Plot authors and publications

Plot average citations vs authors and publications

Explore authors



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



# Mining scholarly relations with Klink

- Klink takes as input a corpus of publications, annotated with keywords
  - Keywords can be user generated or can be automatically extracted from the abstract or the full text of the publication
  - In our experiments we used a corpus of 15M computer science publications obtained from Microsoft Academic Search
- Tidies up the set of keywords by removing keywords that do not denote a research area – e.g., “case study” or “NeOn Project”.
- Automatically computes three types of semantic relationships between the identified research areas.
- Returns a KB of semantic relationships between research areas



# Relations mined by Klink

- *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.
  - 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  $B$ .
  - E.g., Research in “Ontology Engineering” contributes to research in “Semantic Web”

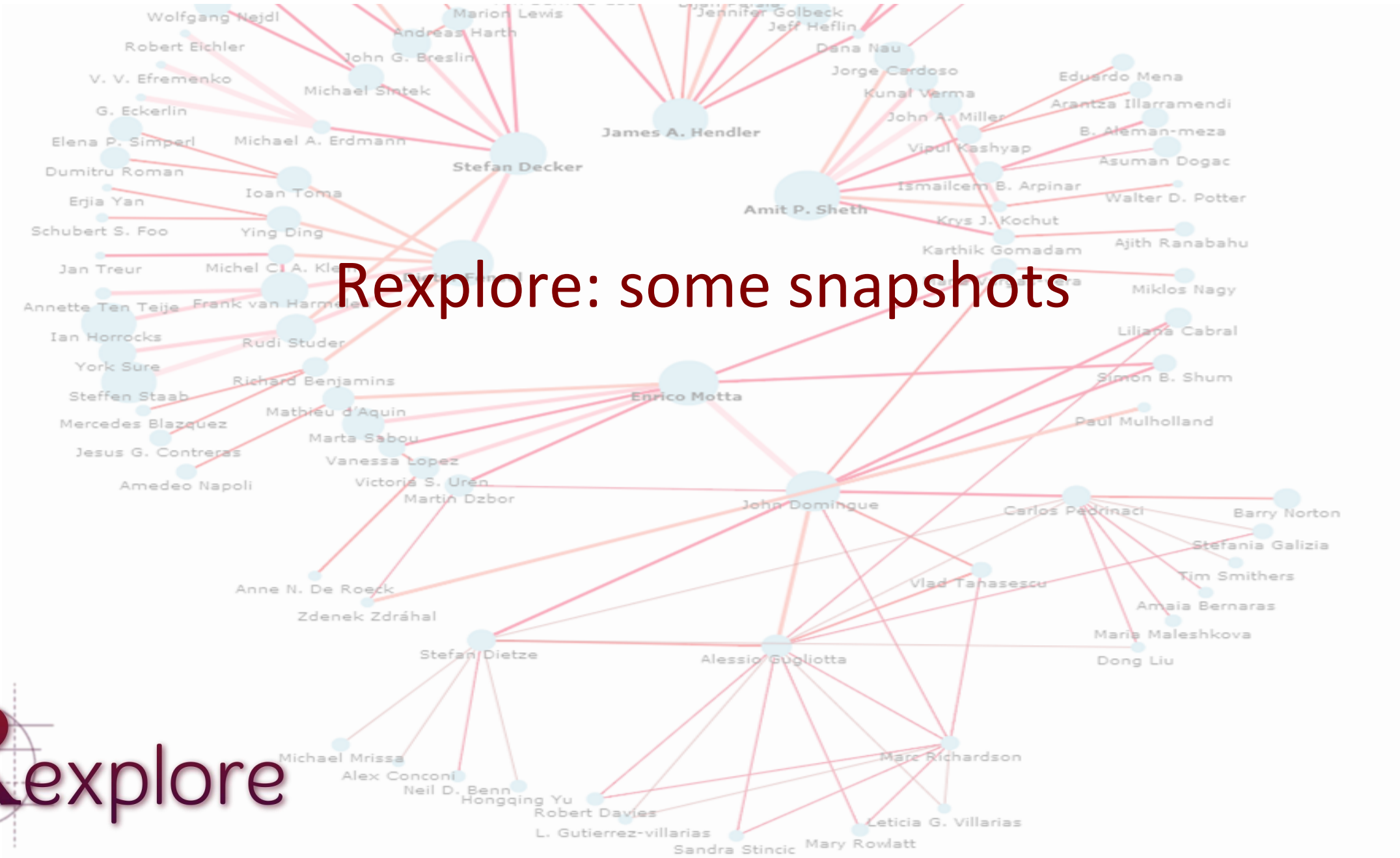
## Semantic Relationships -

### ⊗ 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 +

*From a corpus of 15M papers accessed through the MAS API Klink identified about 1500 research topics and structured them by means of almost 3000 semantic relationships*







**R**explore

# Expert Search (1a)

Authors | Publications

Search Options:  

Author Name

Career Years:  All Years  Select Range

5 to 15

Semantic Web AND Social Network

Turn venue suggestion off - reload

WWW  WI  ISWC  ESWC  OTM Workshops  ASONAM  
 HICSS  CHI  ICWS  CIKM  ICALT  DEXA  SKG  
 SAC  CSE  AAAI  KDD  ICDE  ICWE  KES  ICEIS  
 ASWC  EKAW  Fuzzy Systems and Knowledge Discovery  
 ICDM

All  At least one

Rank result by

Use harmonic mean for multi-dimensional searches

Showing authors 1 - 50 of 81 total results.



1. **Jennifer Golbeck**, University of Maryland (US) - Debut: 2002 -  main node in graph view  
Publications in Semantic Web: **49** - Citations in Semantic Web: **718**  
Publications in Social Network: **44** - Citations in Social Network: **920**  
Total Publications in the topics: 93 - Total Citations in the topics: 1 638  
HM Publications in the topics: **46** - HM Citations in the topics: **807**  
Total Publications: 96 - Total Citations: 1322  
H-Index: 20 - G-Index: 35
2. **Peter Mika**, Yahoo Research Labs (US) - Debut: 2000 -  main node in graph view  
Publications in Semantic Web: **33** - Citations in Semantic Web: **547**  
Publications in Social Network: **9** - Citations in Social Network: **609**  
Total Publications in the topics: 42 - Total Citations in the topics: 1 156  
HM Publications in the topics: **14** - HM Citations in the topics: **576**  
Total Publications: 71 - Total Citations: 1080  
H-Index: 11 - G-Index: 32
3. **Bijan Parsia**, University of Manchester (GB) - Debut: 2001 -  main node in graph view  
Publications in Semantic Web: **57** - Citations in Semantic Web: **1 867**  
Publications in Social Network: **2** - Citations in Social Network: **213**  
Total Publications in the topics: 59 - Total Citations in the topics: 2 080  
HM Publications in the topics: **4** - HM Citations in the topics: **382**  
Total Publications: 141 - Total Citations: 3464  
H-Index: 30 - G-Index: 57
4. **Harith Alani**, University of Southampton (GB) - Debut: 2000 -  main node in graph view  
Publications in Semantic Web: **32** - Citations in Semantic Web: **309**  
Publications in Social Network: **10** - Citations in Social Network: **107**  
Total Publications in the topics: 42 - Total Citations in the topics: 416  
HM Publications in the topics: **15** - HM Citations in the topics: **159**  
Total Publications: 75 - Total Citations: 1043  
H-Index: 18 - G-Index: 30
5. **Boanerges Aleman-meza**, Rice University (US) - Debut: 2003 -  main node in graph view  
Publications in Semantic Web: **24** - Citations in Semantic Web: **451**  
Publications in Social Network: **8** - Citations in Social Network: **73**  
Total Publications in the topics: 32 - Total Citations in the topics: 524  
HM Publications in the topics: **12** - HM Citations in the topics: **126**

*Researchers in the 5-15 career range with expertise in both semantic web and social networks, with publications in at least one of {CHI, ISWC, WWW), ranked with respect to the impact of their work in these two areas (using harmonic mean)*




# Expert Search (1b)

Authors Publications

Search Options:  

Link filter with query


Graph Control 


List graph nodes


Connections:


Ranking:


Style:

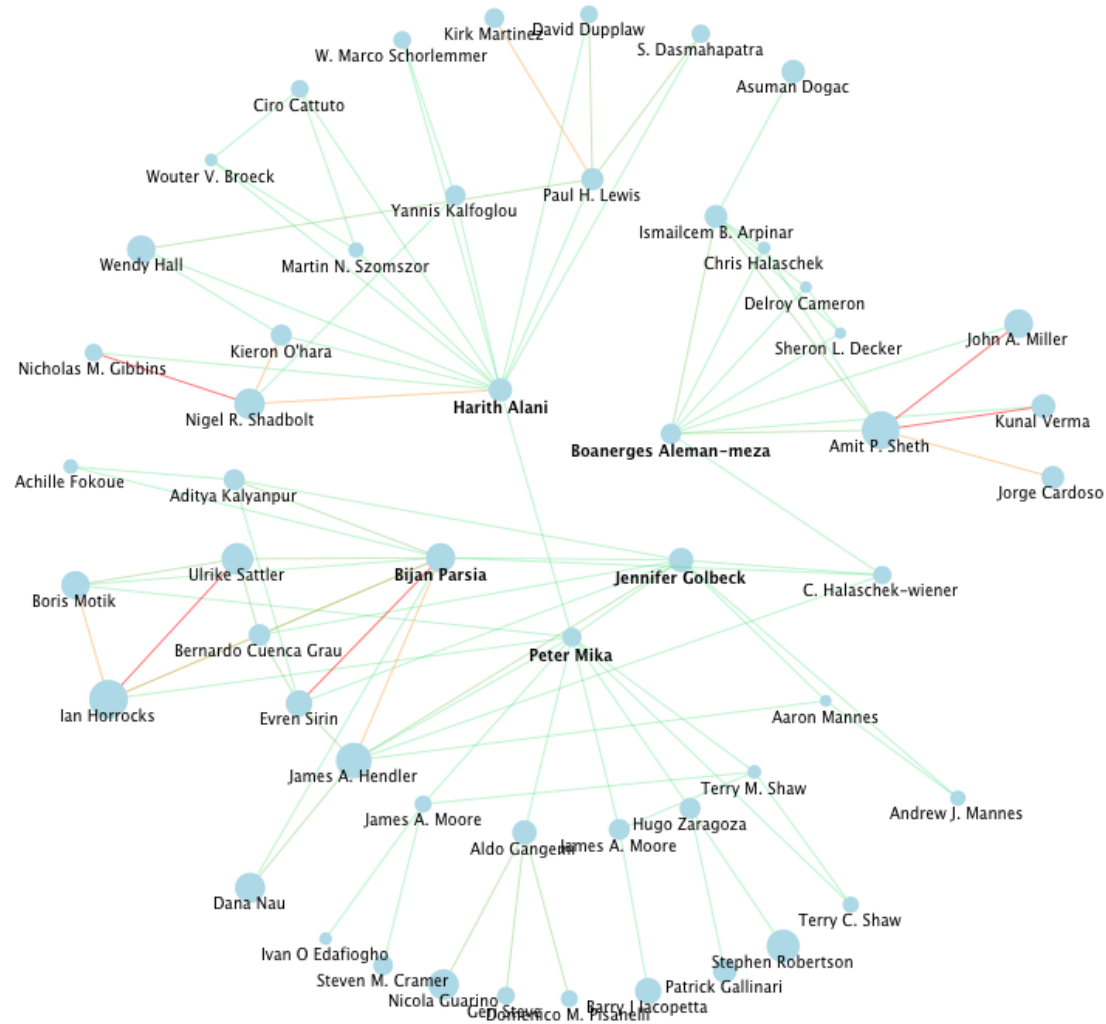
Context 

Filters 

Advanced 

Node info 

Link info 



*Graph view of main researchers identified in previous slide, linking them to their main co-authors.*

*The diameter of a node reflects the h-index of the researcher*

# Expert Search (2)

About Publications Team Contact

Bio Topics Co-authors C.A. Graph View

### Graph Control

Reload Reset List graph nodes

Connections: Co-authors

Ranking: Total Publication in Topic

Style: Standard

### Context

Machine Learning

Suggest venues from topics

Venue

### Filters

Career Years  
 All Years  Select Range 1 to 5

Filter nodes by topic/s in context  
 Filter connections by topic/s in context  
 Filter nodes by venue/s in context  
 Filter connections by venue/s in context

### Advanced


Filtered by:  
Career age: 1 - 5  
Topic (nodes): Machine Learning  
[Turn off filters](#)

The graph displays a network of researchers. The central node is Enrico Motta. He is connected to Gonzalo Melo, Lucia Specia, and Sofia Angeletou. Gonzalo Melo is further connected to Spiros Nikolopoulos, Piercarlo Slavazza, Cristina Kadar, and Sebastian Blohm. Lucia Specia is connected to Zhuoran Wang, Erick Galani Maziero, Volpe Nunes, and Claudio Baldassarre. Sofia Angeletou is connected to Carlo Allocca and Diego De Cao. There are also connections between Spiros Nikolopoulos and Aba-sah Dadzie, and between Piercarlo Slavazza and Cristina Kadar.


*Career-young (1-5) people who have co-authored with Enrico and have expertise in machine learning, ranked in terms of #publications in this topic*


# Shared Research Trajectories


Bio Topics Co-authors C.A. Graph View


Graph Control 


List graph nodes


Connections: Temporal Topic Similarity 


Ranking: H-Index 


Style: Standard 

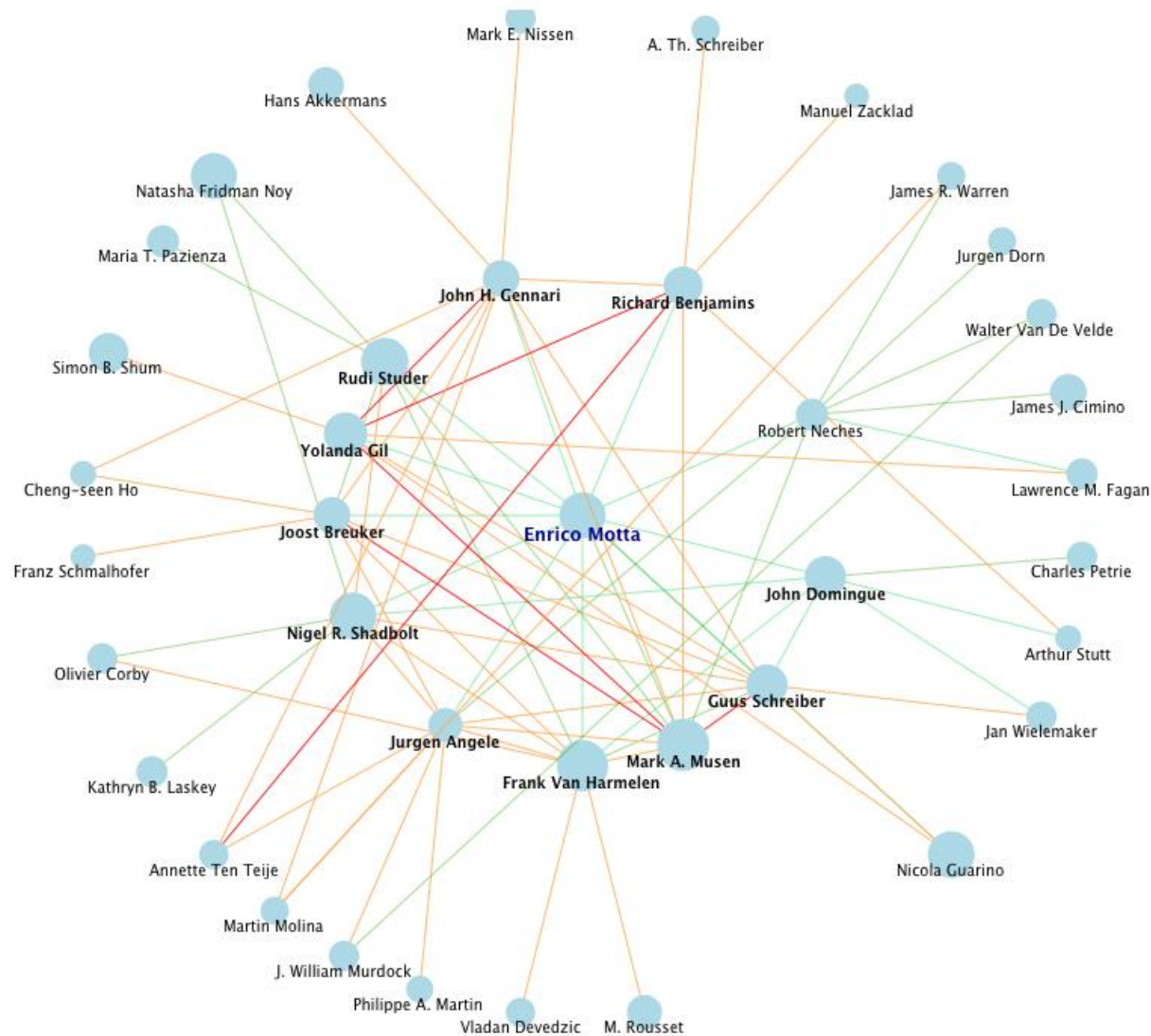
Context 

Filters 

Advanced 

Node info 

Link info 



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

# Normalised impact per topic over time



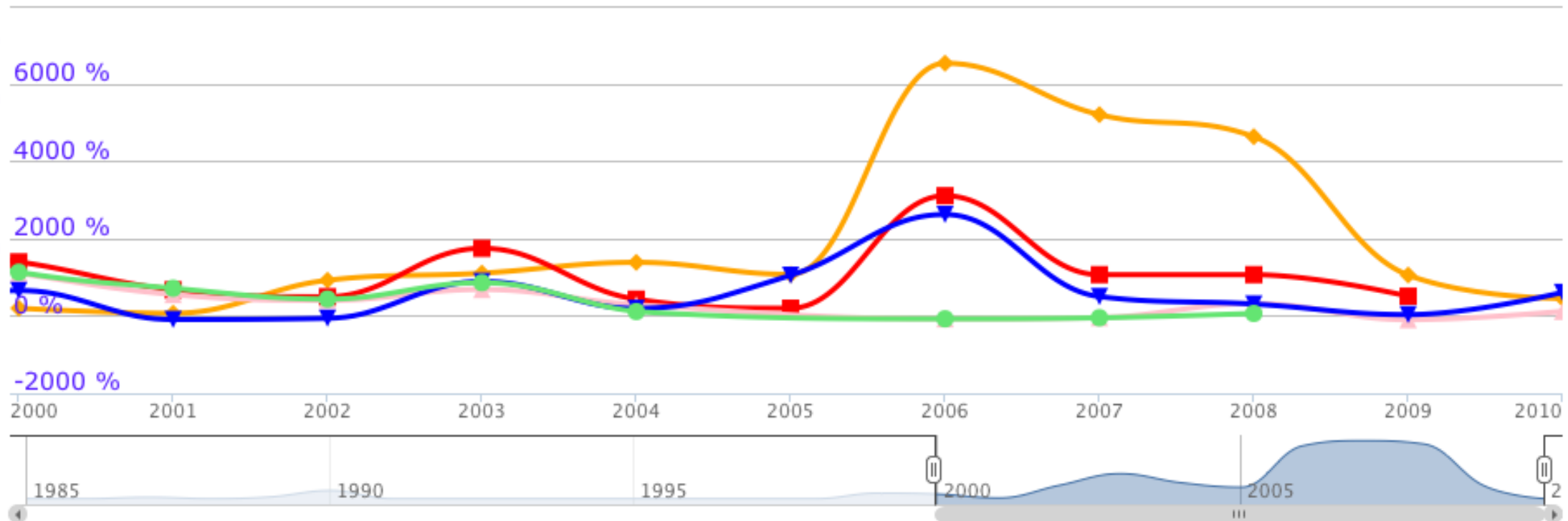
Enrico Motta  
Citations normalized by topics



- total citations
- Semantic Web
- Knowledge Base
- Artificial Intelligence
- Information Retrieval
- Expert System

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

% Difference against average topic cit.



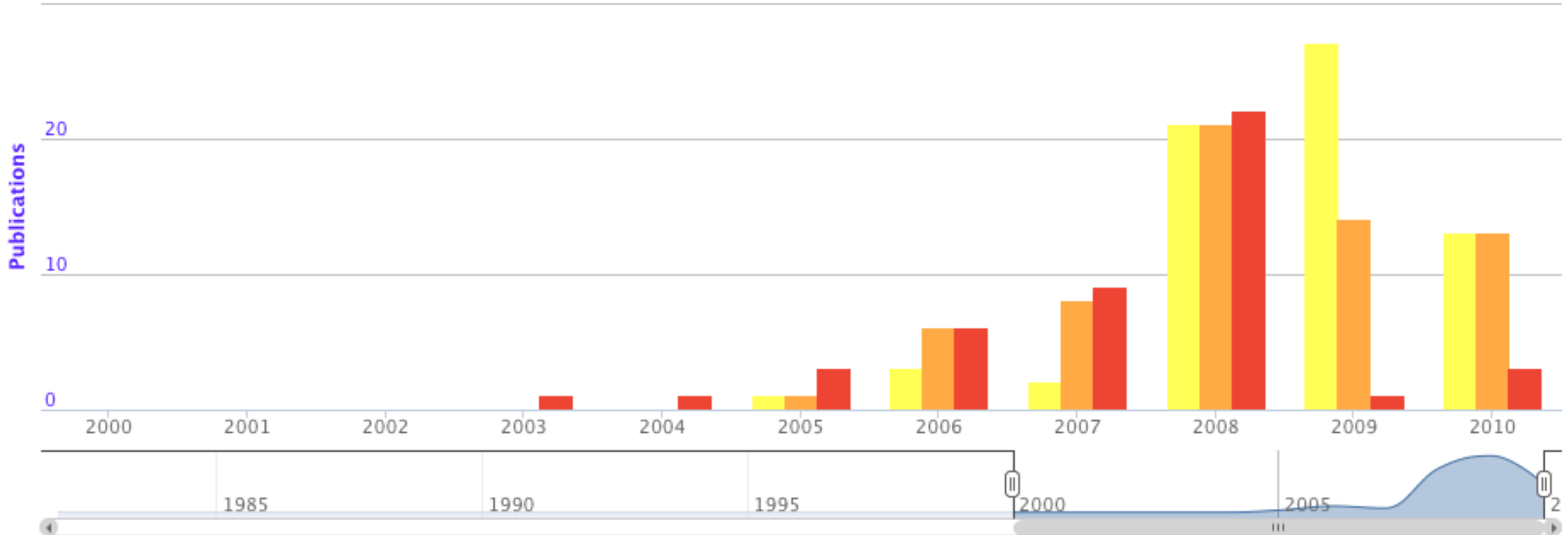
# Where are SW authors going?...

Semantic Web  
Negative Migration Balance



■ Social Network ■ Wireless Sensor Network ■ Service Oriented Architecture

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





# Conclusions (1)

- Rexplore aims to provide an integrated solution to support tasks that require the exploration and analysis of scholarly data
- It does so by integrating a semantic foundation with statistical and visual analytics solutions

# Conclusions (2)



- The fine-grained structure of research topics generated by Klink supports
  - Expert search, trend analysis, and exploration at a very fine grained level of granularity
  - The definition of fine-grained impact metrics, such as “citations in topics” or “normalised impact with respect to topic”, which allow to measure very specific elements of academic impact



# Conclusions (3)



- A rigorous empirical evaluation confirmed:
  - The effectiveness of the functionalities provided by the tool. 94% of the testers described Rexplore as “very effective”
  - The robustness of the tool with respect to tasks proposed by the users themselves. Rexplore was able to support satisfactorily 88% of the testers with respect to tasks proposed by them

KNOWLEDGE MEDIA

K M i

I N S T I T U T E