

Synopsis of Recent and Current Research Activities

October 11, 2015

Yannis Tzitzikas

Information Systems Laboratory (ISL)
Institute of Computer Science (ICS)
Foundation for Research and Technology – Hellas (FORTH)

Outline

- ◆ List of activities
- ◆ One slide per activity
- ◆ Technology offerings



LIST OF ACTIVITIES



List of Activities

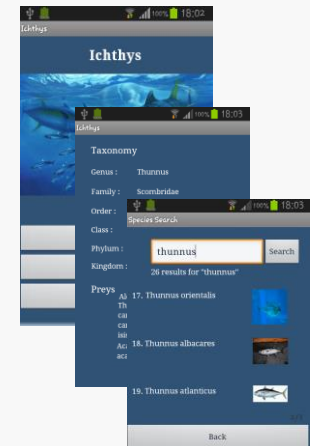
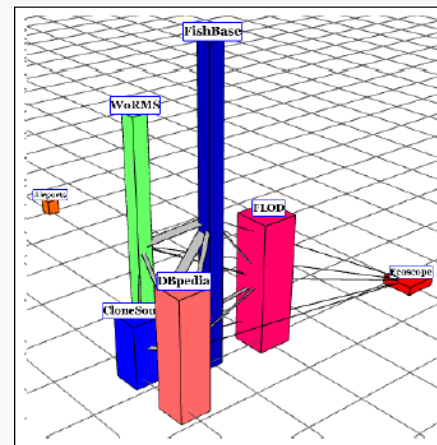
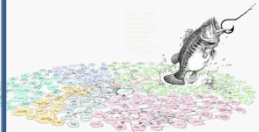
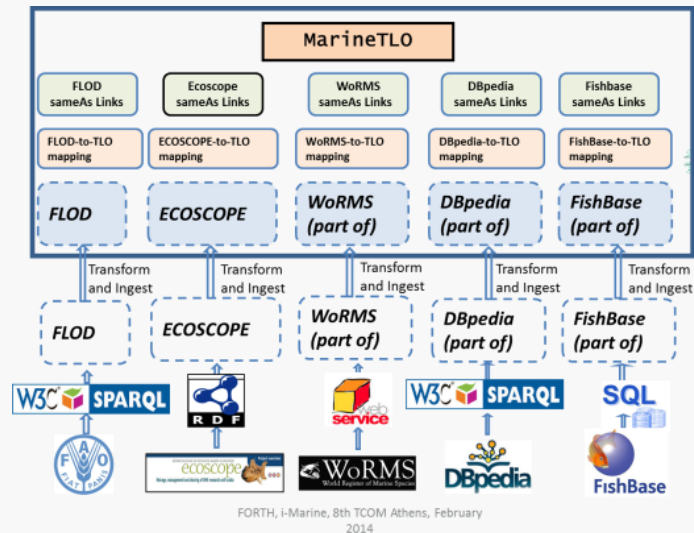
- ◆ **Semantic Integration** of Data
- ◆ **Exploratory** Search
- ◆ Knowledge evolution with emphasis on **comparison**
- ◆ Knowledge Mgmt for **Digital Preservation**



Semantic Integration of Data

MatWare: A framework for **integrating** heterogeneous information and **evaluating their connectivity**

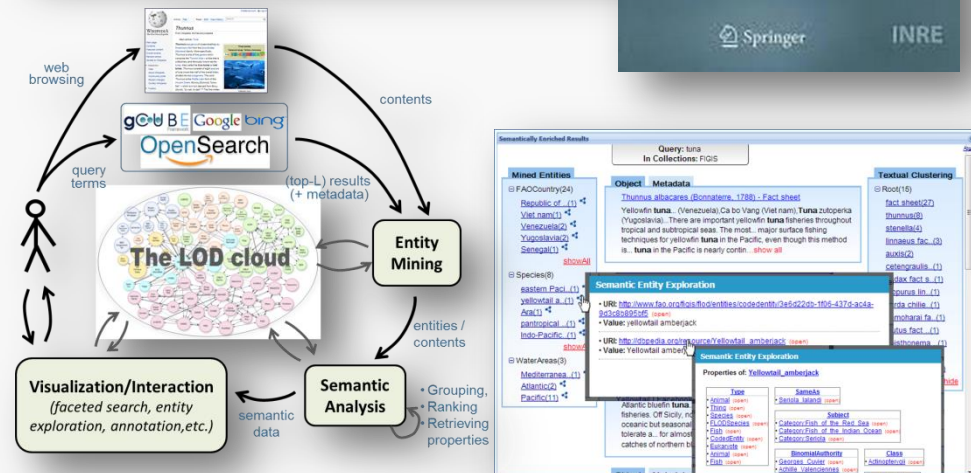
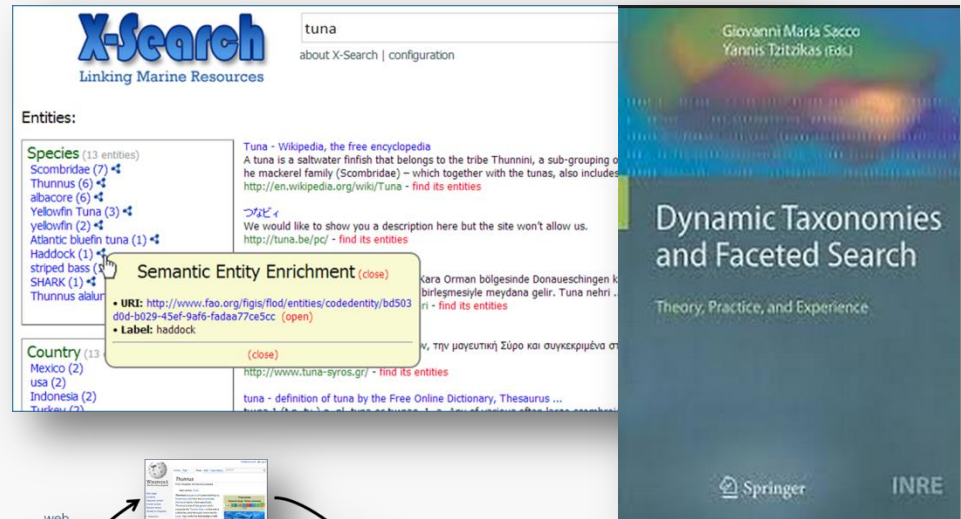
- ◆ Enables the creation of domain-specific semantic warehouses that offer complex query answering
- ◆ Functions: fetching and transformation of data, schema and instance mapping and matching, warehouse refreshing and monitoring.
- ◆ Offers novel metrics (and visualizations) for evaluating the connectivity of the data in the warehouse.
- ◆ Exploited for constructing warehouses for the marine domain
 - ◆ iMarine project, BlueBRIDGE project (ongoing)



Exploratory Search

Research and implementation of novel **exploratory search** services

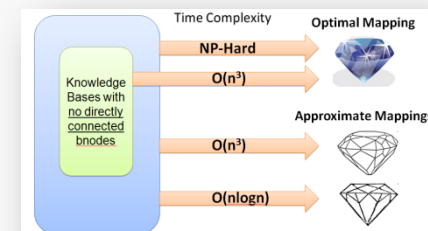
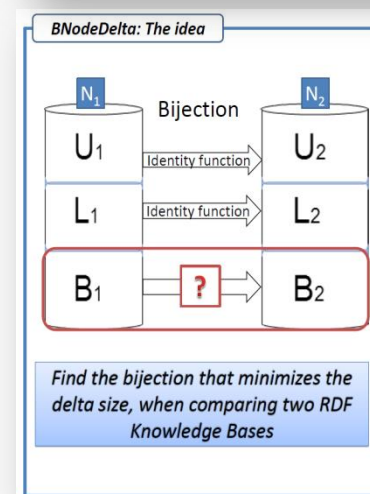
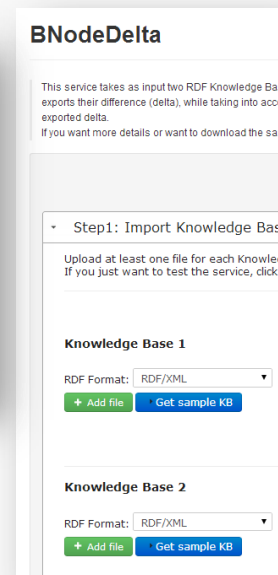
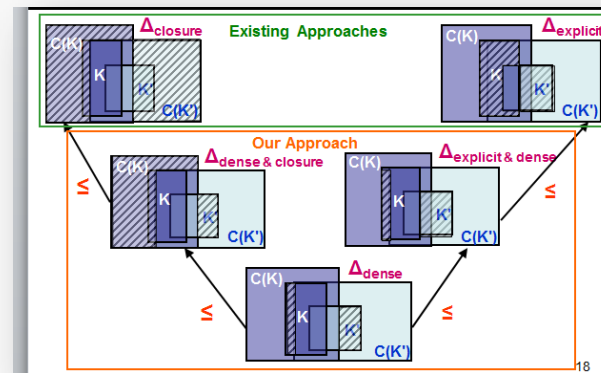
- ◆ Appropriate for recall-oriented information needs (professional search)
- ◆ Applicable over the existing search systems, providing methods that can bridge the gap between the Web of Documents and the Web of Data at Search Time
- ◆ Sophisticated techniques for various steps of the process (autocompletion, semantic analysis of the search results, clustering, preference-based exploration.
- ◆ Implementations in various projects
 - ◆ ESA, gCube e-infrastructure



Knowledge Evolution with Emphasis on Comparison

Contributions in core knowledge evolution techniques with emphasis on **comparisons**.

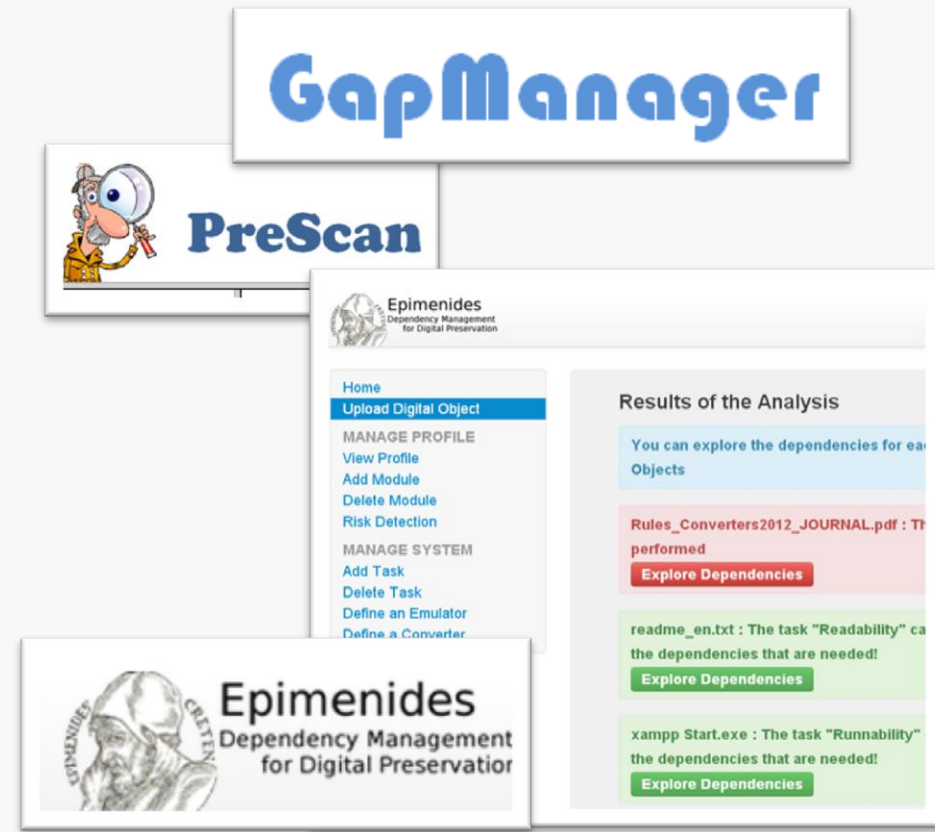
- ◆ The ability to compute the differences that exist between two Knowledge Bases is an important step to cope with the evolving nature of the Semantic Web. RDF Deltas can be employed to aid humans understand the evolution of knowledge, and to reduce the amount of data that need to be exchanged and managed over the network in order to build Semantic Web synchronization, versioning and replication services
- ◆ Various comparison functions
 - ◆ Best paper award at the International Semantic Web Conference 2007
- ◆ Theoretical and practical results concerning the problem of mapping anonymous nodes
 - ◆ Bnode Matching



Knowledge Management for Digital Preservation

Novel ways to exploit **knowledge management** techniques for the needs of **Digital Preservation**

- ◆ Novel knowledge-based Dependency management techniques for the needs of Digital Preservation
- ◆ Development of prototypes and systems that realize the research results
 - ◆ GapManager, PreScan, Epimenides
- ◆ Participation to the main EU projects on Digital Preservation
 - ◆ CASPAR, SCIDIP-ES, APARSEN NoE



TECHNOLOGY OFFERINGS



List of technology offering

◆ Exploratory (Web) Searching

- ◆ **X-Search**: Semantic Analysis of Search Results. First release: 2013 ([leaflet](#))
- ◆ **Hippalus**: Preference-Enriched Faceted Exploration. First release: 2013 ([leaflet](#))
- ◆ **IOS**: A family of novel services for instant exploratory search. First release: 2012
- ◆ **StellaSearch**: A Meta Search Engine over Google offering real time results clustering. First release: 2009
- ◆ **FleXplorer**: A general purpose API for supporting exploratory search. First release: 2009

◆ Semantic Data Management

- ◆ **MatWare**: Tool for automating the creation of Semantic Warehouses. First release 2014 ([leaflet](#))
- ◆ **Bnodeland (the land of blank nodes)**: Compare RDF Graphs with Bnode Matching! First release 2012. Also a generator of synthetic data. First release 2014
- ◆ **RIMQA (RDF Instance Migration Quality Assistant)**: A tool for quality management after schema migration. First release: 2012.
- ◆ **StarLion**: A tool for visualizing RDFS ontologies (with emphasis on top-K diagrams and on exploration services). First release: 2008

◆ Digital Preservation (tools and systems)

- ◆ **Epimenidis**: Automated Reasoning for Digital Preservation. First release 2013 ([leaflet](#))
- ◆ **PreScan**: Automated extraction of file-embedded metadata from file systems. First release: 2008
- ◆ **GapManager**: A dependency management system for the preservation of digital objects. First release: 2008

◆ Ontologies

- ◆ **MarineTLO**: A top-level ontology for the marine domain. First release 2012. ([leaflet](#))



Thanks for your attention

◆ Contact information:

- ⊕ Yannis Tzitzikas
- ⊕ tzitzik@ics.forth.gr
- ⊕ www.ics.forth.gr/~tzitzik

