Wimmics - Olivier Corby

Web-Instrumented Man Machine Interactions, Communities and Semantics
Wimmics (ex Edelweiss)

- 2007 Edelweiss creation (after Acacia)
  Rose Dieng-Kuntz scientific leader
- 2008 Olivier Corby interim leader, Fabien Gandon deputy leader
- 2009 Research topics restructured
- 2011 End of Edelweiss
- 2012 Wimmics: follow-up team, Fabien Gandon team leader
Edelweiss topics

Knowledge Engineering for the Social Semantic Web

2. Graph based Knowledge Representation
3. Interaction Design for the Web

- Knowledge Engineering : Ontologies & Folksonomies
- Information Retrieval & Knowledge Sharing on the Web
- Annotation of Resources (Anything with an URI)
- Human-Computer(-Human) Interaction
- Labeled Graph Match, Social Network Analysis
Edelweiss topics

Methods & Models
Knowledge Engineering, Ergonomics, Semantic SNA, Graph matching

Tools
Corese, Sewese, Ecco, Semantic Wiki, Isicil platform

Applications
Bio Medicine, Geology, eLearning, Corporate Intelligence, Engineering Design
Edelweiss Team October 2011

1. Olivier Corby, CR INRIA
2. Fabien Gandon, CR INRIA, HDR
3. Alain Giboin, CR INRIA
4. Christine Foggia Admin,
5. Erwan Demairy, Eng. INRIA,
6. Nicolas Delaforge, Eng. ANR
7. Michel Buffa, CE UNS
8. Catherine Faron-Zucker, CE UNS
9. Isabelle Mirbel, CE UNS, HdR
10. Adrien Basse, PhD, UGB
11. Luca Costabello, PhD, INRIA
12. Corentin Follenfant, PhD SAP
13. Maxime Lefrançois, PhD, UNS
14. Nicolas Marie, PhD ALU
15. Rakebul Hasan, PhD ANR
16. Oumy Seye PhD, UGB
17. Elena Cabrio, PostDoc
18. Serena Villata, PostDoc
Semantic Annotation of Information Resources

Improve Information Retrieval using Ontology based Semantic Annotations

Extract Semantic Annotation from NL Text by semi automatic methods
Semantic Annotation of Information Resources

- Natural Language text mining
- GATE NLP Platform and Ontology with Terminology
- Identification of Instances and Relations
- Generation of RDF annotations
- Contextual Annotations using RDF Named Graphs
- Ambiguity solving using semantic distance
Graph based Knowledge Representation

- Knowledge Engineering with Graph based languages
- From Conceptual Graphs to RDF (standard in KE)
- Semantic Web, Linked Open Data, Social Network Analysis, Semantic Annotation, Ontology
- Information Retrieval in Labeled Graph
- Graph Match Algorithms
- SPARQL Query Language
SPARQL Extension: Property Path  Edge enumeration

Path variable:

?x foaf:knows+ :: $path ?y

Graph pattern overload:

graph $path { ?a foaf:knows ?b }
SPARQL Extension: Count Nodes

```
select ?x ?y (count(distinct ?a) as ?c)
where {
  ?x foaf:knows+ :: $path ?y
  graph $path {?a foaf:knows ?b filter(?a != ?x)}
}
group by ?x ?y
```

How many intermediate nodes between x and y?
SPARQL Extension: count paths

select ?x ?y ?a (count($path) as ?c)
where {
  ?x foaf:knows+ :: $path ?y
  graph $path {?a foaf:knows ?b filter(?a != ?x)}
}
group by ?x ?y ?a

How many x-to-y paths are passing by a?
SPARQL Extension: Shortest Path

```
select ?x ?y (min (pathLength($path) as ?min)
where {
   ?x foaf:knows+ :: $path ?y
}
  group by ?x ?y
```

Shortest path length between x and y?
Corese Semantic Web Factory

- RDF RDFS XSD
- SPARQL 1.1 Query & Update
- W3C Test cases
- Generic SPARQL Interpreter on Labeled Graphs
- 3 Ports: Corese, Kgraph, Jena
- SPARQL Extensions: Approximate Search wrt RDFS Classes, sql(), xpath()
- Construct-where inference rules
- Pipeline: Query, Rule, if-then-else, Pipeline
- Event Listener
Corese Semantic Web Factory

Open Source, CeCILL-C, INRIA forge
45 applications, 21 PhD Thesis
Teaching: 8 universities
Projects:
  - EU: Sealife, Palette, SevenPro
  - ANR: Neurolog, eWOK Hub, Isicil, Kolflow
Users: IGN, CSTB, IFP, EADS, BRGM, Ademe, INRA, I3S, etc.
Interaction Design of SSW Applications

Aim

Designing Social Semantic Web (SSW) applications adapted to communities

SSW applications = Applications based on a combination of Semantic Web & Web 2.0 (or Social Web) technologies
Interaction Design of SSW Applications: Challenges

- **Reconciling Semantic Web and Social Web approaches to the design of the Web**

<table>
<thead>
<tr>
<th></th>
<th>Semantic Web</th>
<th>Social Web</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>User participation</strong></td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td><strong>Formality</strong></td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td><strong>Inferential capabilities</strong></td>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

- **Articulating developers’ and users’ representations and operating modes**
  - Making the representations “interoperable”

<table>
<thead>
<tr>
<th></th>
<th>Developers</th>
<th>Users</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Representations</strong></td>
<td>Formal</td>
<td>Informal</td>
</tr>
<tr>
<td><strong>Operating modes</strong></td>
<td>Logics of functioning</td>
<td>Logics of use</td>
</tr>
</tbody>
</table>
Interaction Design of SSW Applications: Solutions explored

Reconciling Semantic Web and Web 2.0 approaches
Articulating developers’ and users’ representations/operating modes

Development of participatory design methods
adapted to the design of SSW applications

Development of collaborative and participatory tools
for supporting the design of SSW applications

Modeling of “collectives” (communities, groups, etc.),
of their members and of their interactions

Recognizing/Visualizing “collectives”
Cf. Visualizing networks and graphs
SELECTED APPLICATIONS
Interaction Design of SSW Applications: Participatory ontology engineering method

• **Adaptation of a method** designed in the Acacia project

• **Requirements**
  • Making the method collaborative
  • Allowing the participation of end-users
  • Designing tools/functionalities to support the application of the collaborative method

• **Method developed**
  • in the context of: European project PALETTE / ANR project e-WoK_HUB
  • in connection with the development of the ECCO collaborative ontology editor
Ontology Engineering Method

- Collaborative Workflow
- Distribution of tasks between Developers and Users

*Context: European project PALETTE, ANR project e-WoK_HUB*
Interaction Design of SSW Applications: 

**ECCO — A Collaborative and Contextual Ontology Editor**

**Involving users in ontology editing**

**Making users and developers collaborate**

**Context**: ANR project **e-WoK_HUB**, European project **PALETTE**

---

**ECCO Designer - Editeur d’ontologies**

**Manual extraction of terms**

**Automatic extraction of terms**

---

**Discussion à propos du terme "GeochronologicUnit"**

- Laura Masse : We decided to represent dates as instances (and not as concepts) because of a particularity of the timescale, we can find alternative boundaries to a single unit.
- Laura Masse : We decided for using a standard ontology model (OWL) rather than the one modeled in the original paper.
- Francesco : The mapping with the timescale: the International Standard Scale and the regional version, with the terms used in Europe. The geologic units are going to be represented as instances of the class GeochronologicUnit of the QTS ontology (in fact, if its sub-concepts Eon, Era, etc.). For example, Jurassic is an instance of the concept Period, and Eocene is an instance of the concept Creta.
- Laura Masse : In the original proposal, the kinds of units are represented just as a textual attribute. Here these are represented as concepts (Eon, Era, Period, etc.).
Assisted Structuration of Folksonomies

web 2.0

Flat folksonomy

thesaurus

polluant

énergie

related

related

pollution

has narrower

pollutions du sol
Combining metrics

Orthographic metric
Monge-Elkan Soundex, JaroWinkler, asymmetric Monge-Elkan Qgram

Contextual metric
cosine vectors co-occurrent tags

Social metric
include user communities
Interaction Design of SSW Applications: Tag Editor/Browser

Involving users in tags structuring

Context: ANR project ISICIL

Automatically suggested semantics

- environnement- pollution
- pollution atmospherique impact sante
- traceur de pollution
- halo de pollution
- pollution development instrumental

Search: pollution

pollution transfrontaliere
pollution visuelle
pollution organique
pollution harmonique
pollution photooxydante
pollution episodique
Social Network Semantic Analysis

$\delta^\circ (p) = \{ x; \text{rel}(x, p) \}$

$$d^\circ_{in}(p) = \{ x; \text{rel}(x, p) \}$$

Social network analysis

Semantic web is not antisocial
ex. qualified degree

degree(x) = 5

degree <famille>(x) = 3
## Formal definition in SPARQL

<table>
<thead>
<tr>
<th>SNA indices</th>
<th>SPARQL formal definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Comp}_{rel&gt;(G)}</td>
<td><code>select ?x ?y from &lt;G&gt; where { ?x param[rel] ?y } group by any</code></td>
</tr>
<tr>
<td>\textit{D}_{rel,dist&gt;(y)}</td>
<td><code>select ?y count(?x) as ?degree where { {?x (param[rel])*::$path ?y filter(pathLength($path) &lt;= param[dist]) UNION {?y param[rel]::$path ?x filter(pathLength($path) &lt;= param[dist])}} group by ?y</code></td>
</tr>
<tr>
<td>\textit{D}_{in} \textit{rel,dist&gt;(y)}</td>
<td><code>select ?y count(?x) as ?indegree where {?x (param[rel])*::$path ?y filter(pathLength($path) &lt;= param[dist])} group by ?y</code></td>
</tr>
<tr>
<td>\textit{Diam}_{rel&gt;(G)}</td>
<td><code>select pathLength($path) as ?length from &lt;G&gt; where { ?y s (param[rel])*::$path ?to } order by desc(?length) limit 1</code></td>
</tr>
<tr>
<td>\textit{nb}_{rel&gt;(from,to)}</td>
<td><code>select ?from ?to count($path) as ?count where {?from sa (param[rel])*::$path ?to } group by ?from ?to</code></td>
</tr>
<tr>
<td>\textit{nb}_{rel&gt;(b,from,to)}</td>
<td><code>select ?from ?to ?b count($path) as ?count where {?from sa (param[rel])*::$path ?to graph $path(?b param[rel] ?j) filter(?from != ?b) optional { ?from param[rel]::$p ?to } filter(!bound($p)) } group by ?from ?to ?b</code></td>
</tr>
<tr>
<td>\textit{C^c}_{rel&gt;(y)}</td>
<td><code>select distinct ?y ?to pathLength($path) as ?length (1/sum(?length)) as ?centrality where { ?y s (param[rel])*::$path ?to } group by ?y</code></td>
</tr>
</tbody>
</table>

**CORESE/KGRAM**
Recognizing and visualizing communities in social networks

Context: ANR project ISICIL, INRIA ADT SegViz (Gephi Plugin)

1. pollution
2. sustainable development
3. energy
4. chemistry
5. air pollution
6. metals
7. biomass
8. wastes

Mixed graph of actors and interests
Selected Publications

- Corby et al. The KGRAM Abstract Machine for Knowledge Graph Querying. *IEEE/WIC/ACM* 2010
Completed Projects

**SeaLife** EU: SW in Bio Medicine

**SevenPro** EU: SW in Engineering Design

**Palette** EU: SW in Education

**e-WOK Hub** ANR: SW in Geology

**ImmunoSearch, Biomarker**: SW in Genetics

**Griwes** Color INRIA: Generic Graph Model

**DESIR** Color INRIA: with INRA & UNS
Collaborations

Alcatel Lucent PhD Thesis
SAP PhD Thesis
IGN Master Thesis
Orange Labs, Atos, Mondeca, EADS : ANR Projects
Fraunhofer, London U., EPFL, Biotec Dresden, etc.
EU Projects
Collaborations

Kewi I3S CNRS UNS
Orpailleur, Exmo, Silex LIRIS, GDD Nantes, Tatoo, RCR/GraphIK, Tech Cico UTT, Telecom ParisTech, Eurecom

INSEE, Ademe, Fing, BRGM, IFP, ENSMP, LISI
Standardization

W3C WG
• GRDDL
• RDFa
• SPARQL 1.1
• RDF 1.1
• Provenance
Education

• UNS Semantic Web & Knowledge Engineering, Master & Licence Pro
• UNS GUI, Ergonomics of IT, Master
• UGB Senegal Semantic Web
• Ecole Centrale Semantic Web Introduction
• SKEMA/Ceram Semantic Web Introduction
Misc.

• co Chair of WWW 2012
• co Chair of *Social Web & Web Science* at ESWC 2012
• Will represent INRIA at W3C
• Contint Program Committee
• GDR Psycho Ergo, member of Concil
• Organizing Committee of EPIQUE 2009
• Chair IC 2009
• dbpedia.fr
• ILab in preparation
• Book on Semantic Web with Dunod
Web-Instrumented Man-Machine Interactions, Communities, and Semantics*

a proposal for a joint research team between INRIA Sophia Antipolis –Méditerranée and I3S (CNRS and University of Nice – Sophia Antipolis).

(*) wimmics comes from wimi, a variety of roses.
**members**

**Head (and INRIA contact):** Fabien Gandon

**Researchers**
1. Michel Buffa, MdC (UNS)
2. Olivier Corby, CR1 (INRIA)
3. Catherine Faron-Zucker, MdC (UNS)
4. Fabien Gandon, CR1, HDR (INRIA)
5. Alain Giboin, CR1 (INRIA)
6. Nhan Le Thanh, Pr. (UNS)
7. Isabelle Mirbel, MdC, HDR (UNS)
8. Peter Sander, Pr. (UNS)

**Post-doc**
1. Elena Cabrio (INRIA)
2. Serena Villata (ANR Datalift)

**Research engineers**
1. Julien Cojan (INRIA)
2. Nicolas Delaforge (INRIA, ANR ISICIL)
3. Erwan Demairy (INRIA, ADT)

**Assistants**
Christine Foggia (INRIA) and Marie Hélène Prosillico (I3S)

**PhD students**
1. Pavel Arapov, 1st year (EDSTIC-INRIA)
2. Adrien Basse, 2nd year (UGB-INRIA)
3. Franck Berthelont, 2nd year (UNS-EDSTIC)
4. Ahlem Bouchahda, 3rd year (UNS-SupCom Tunis)
5. Khalil Riad Bouzidi, 2nd year (UNS-CSTB)
6. Luca Costabbello, 1st year (INRIA-CORDI)
7. Corentin Follenfant, 1st year (SAP)
8. Maxime Lefrançois, 1st year (EDSTIC-INRIA)
9. Nicolas Marie, 1st year (Bell-ALU, INRIA)
10. Hasan Rakebul, 1st year (ANR-Kolflow)
11. Oumy Seye, 1st year, (Rose Dieng allocation)
12. Imen Tayari, 3rd year (UNS-Sfax Tunisie)
13. Viet-Hoang Vu, 4th year (UNS-Factory)
graphs, graphs…

Ipernity social network structure

Interest graph

Mixed graph of actors and interest

Folksonomy restructuring

Interest matrix

Clusters in KmA
challenges
analyzing, modeling, formalizing and implementing graph-based social semantic web applications for communities.

multidisciplinary approach for analyzing and modeling
- the many aspects of intertwined information systems
- communities of users and their interactions

formalize and reason on these models
- new analysis tools and indicators
- new functionalities and better community management.
interactions

- analyzing & modeling communities
- interacting through social semantic web app.
- interacting with dynamic semantic web app.

how do we improve our interactions with such an information system that keeps getting more and more complex?

how do we reconcile and integrate the formalized stable semantics of computer science and the negotiable social interactions?

how do we reconcile local contexts of users and global characteristics of the world-wide virtual machine and information systems that the web has become?
typed graphs

- formalizing models and implementing social semantic web applications
- calculating on heterogeneous joined typed graphs of the web

what kind of formalism is the best suited for such models?
how do we analyze these typed graph structures and their interactions?
how do we support different graph life-cycles, calculations and characteristics in a coherent and understandable way?
ANR projects

**isicil.inria.fr (2/3)**
- enterprise social networking
- business intelligence, watching, monitoring
- communities of interest, of practice, of experts

**datalift.org (1/3)**
- frow raw public data to interlinked data and schemas
- a platform and documentation to assist the process
- validation on real datasets

**kolflow.univ-nantes.fr (1/3)**
- reduce the overhead of communities in the process of continuously building knowledge
- federated semantic wikis as a distributed blackboard for man-machine collaboration
projects

dbpedia.fr
• French version of dbpedia from wikipedia
• Ministry of culture

SeGViz
• Semantic Graph Visualization
• INRIA Grant for software development
Thank you for your attention