



# 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

1. Semantic Annotation of Information Resources (2007-2009)
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 {
```

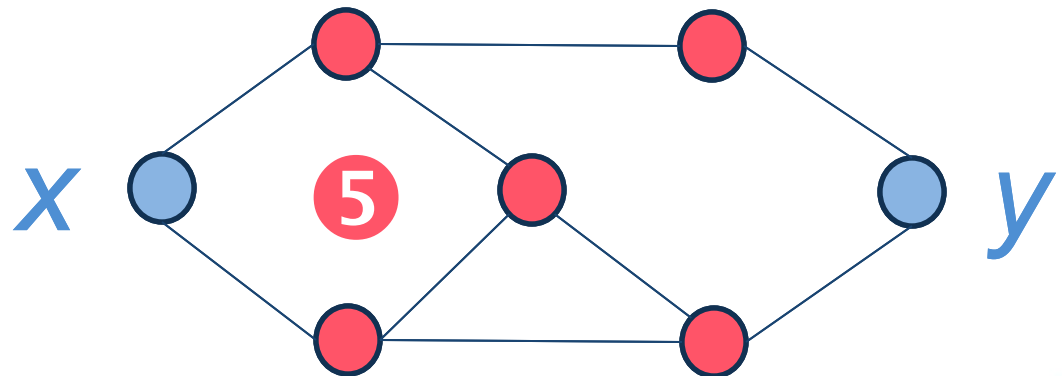
```
  ?x foaf:knows+ :: $path ?y
```

```
  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 {
```

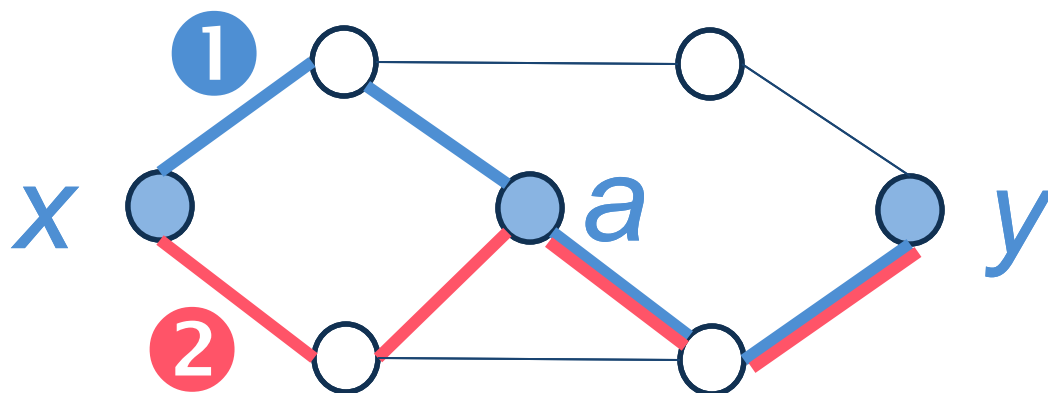
```
  ?x foaf:knows+ :: $path ?y
```

```
  graph $path {?a foaf:knows ?b filter(?a != ?x)}
```

```
}
```

```
group by ?x ?y ?a
```

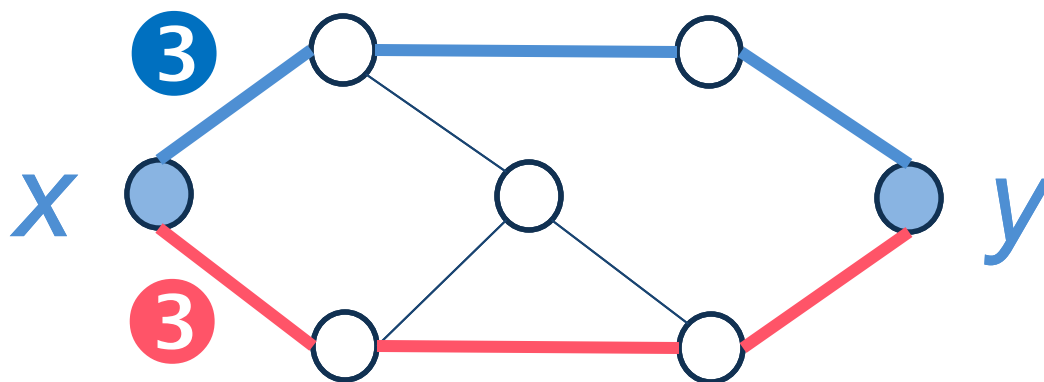
*How many x-to-y paths  
are passing by a?*



## 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**

	Semantic Web	Social Web
<i>User participation</i>	Low	High
<i>Formality</i>	High	Low
<i>Inferential capabilities</i>	High	Low

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

	Developers	Users
<i>Representations</i>	Formal	Informal
<i>Operating modes</i>	Logics of functioning	Logics of use



# 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

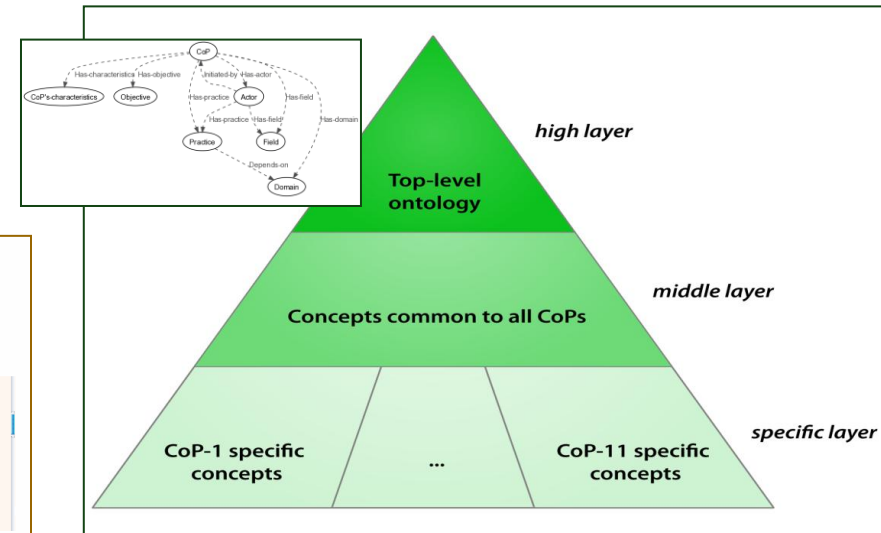
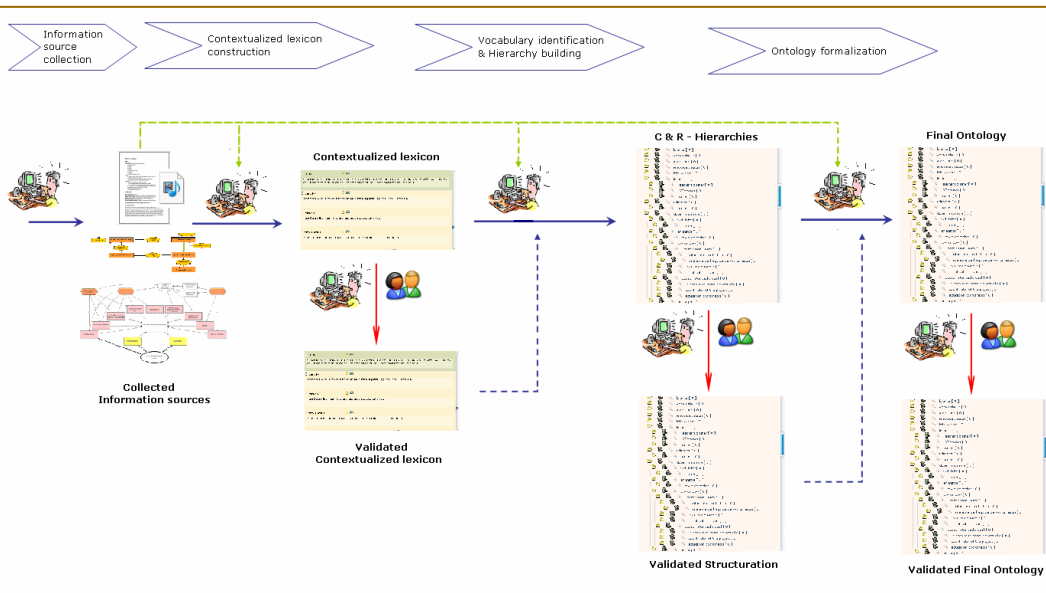
# SELECTED APPLICATIONS

## 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



CoP-specific Ontologies	CoPs
<i>O' ICTE</i>	TIC-EF/TIC-FA
<i>DocHETICE</i>	Form@HETICE
<i>O'Learn-Nett</i>	Learn-Nett
<i>Technical, Human, Learning/Training Problems</i>	@pretic
<i>"Transition Formation Travail" Ontology</i>	TFT
<i>"Pedagogical Resource Management"</i>	CoPe-L
<i>"WikiPrépas"</i>	ePreP

- 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 - ÉDITEUR D'ONTOLOGIES  
de la conception à l'évolution...

Données → Extraits-Termes → **Vocabulaire** → Hiérarchie → Ontologie

Marquer comme: Nouveau synonyme Nouvelle définition

isCorrelatedWith

Eonothem

GeochronologicUnit

Erathem

hasPreviousUnit

StratigraphicBoundary

hasEndDefinedBy

hasStartDefinedBy

Aucune discussion existante

Discussion à propos du terme "GeochronologicUnit"

Laura Mastella : We decided to represent units as instances (and not as concepts) because of a particularity of the time scale: we can find alternative boundaries to a same unit.

Laura Mastella : We associated this ontology with the OWL Time Ontology: <http://www.isi.edu/~pan/OWL-Time.html>

Laura Mastella : We decided for using a standard ontology model (OWL) rather than the time model used in the original paper.

Laura Mastella : We are working with 2 versions of a Time Scale: the International Standard Scale and the regional version, with the terms most used in Europe. The geological units are going to be represented as instances of the concept GeochronologicUnit of the GTS Ontology (in fact, of its sub-concepts Eon, Era, etc.). For example, Jurassic is an instance of the concept Period, and Cenozoic is an instance of the concept Era.

Laura Mastella : In the original proposal, the kinds of units are represented just as a textual attribute. Here there are represented as concepts (Eon, Era, Period etc)

Auteur

profil de l'auteur (ici: ingénieur)

Données → Extraction de termes → **Vocabulaire** → Hiérarchie → Ontologie

Saisir le nom de l'ontologie concernée\*: U

Sélectionnez, dans la source de données,

Manual extraction of terms

Terms choisis:

- Enseignant
- Dominique Bouillet
- accompagnons les enseignants
- cellule des ressources multimédias pour la pédagogie
- coordinateur du module UX11
- Denis
- Frédéric

Extraction de termes

Mots extraits:

- eau de la nappe (1)
- nappe du dogger (1)
- eau de chlorite (1)
- mole litre (1)
- ment majeur (1)
- tousle autre (1)
- anion majeur (1)
- dolomite ddsordonn (1)
- kaolinite d (1)
- rail tel (1)
- h silice (1)
- température proche de la rdalid (1)
- température proche de la température proche (1)
- température proche de la température (1)
- température proche (1)
- température moyen (1)
- variation local (1)
- rapport c (1)
- origine de chlorit (1)

Automatic extraction of terms

Acabit Auto Extraction Fastr Auto Extraction

Sort by fréquence

mots que j'ai extraits définitions/contextes que j'ai extraits

Savegarder Annuler

# Assisted Structuration of Folksonomies

## web 2.0

**Fabien Gandon's Bookmarks**  
Bookmarks | Network | Tags | Subscriptions | Inbox  
Site: <http://www-sop.inria.fr/acacia/personnel/Fabien.Gandon/>  
See more bookmarks in [Popular](#), [Recent](#), or [look up a URL](#).

gandon Type a tag Bookmarks 276  
Display options

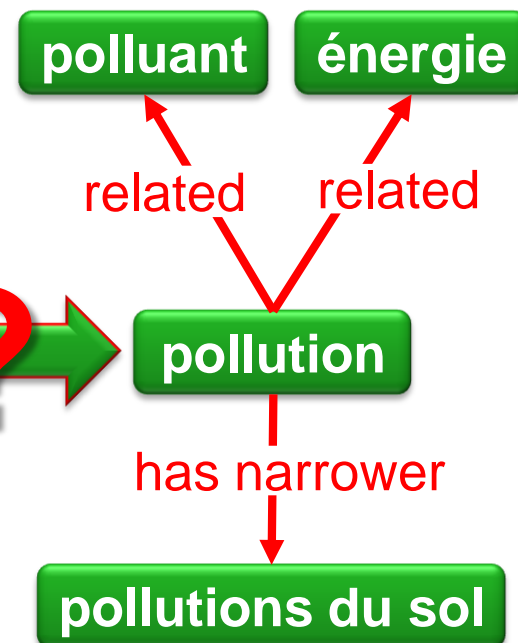
Semweb Pro : le web sémantique français se cherche encore des entreprises utilisatrices::Gestion des données::LeMagIT 2  
A l'occasion de la conférence Semweb Pro qui s'est déroulée les 17 et 18 janvier, la communauté du web sémantique français a fortement insisté sur les usages, histoire de rappeler aux concepteurs que le concept du web sémantique est bel et bien sorti des labos.  
EDIT | SHARE | DELETE semweb semwebpro semanticweb semantic\_web presse media

Semweb.pro : la France bien positionnée sur le web sémantique 5  
Pour cette première édition, la conférence française a réuni la communauté et présenté des cas d'utilisation concrets.  
EDIT | SHARE | DELETE semantic\_web semanticweb semweb France conference semwebpro media presse

## Flat folksonomy

communciation communication compareteur  
compensation comportement condition  
conference confiance consommation  
contact convention cosmetique coton  
covoiturage creation critique CSS cuisine  
dechet dell demarche developpement document  
donnee durable dynamique eau ecoconception  
ecolife ecologie ecomerce economie  
elavage emballage emission empreinte encre  
encyclopedie energie enquete  
entreprise environnement eolien Espagne  
ericsson espee ethique etiquette etude  
europe evaluation evolution extensif faune  
federation fertilisation filiere film finance flore  
fondation forum france fruit GES  
goodguide grandedistribution graphique green  
greenpeace greenwashing grenelle guide  
habitat hitech HQE HTML hybride impact  
import imprimante inde indice industrie

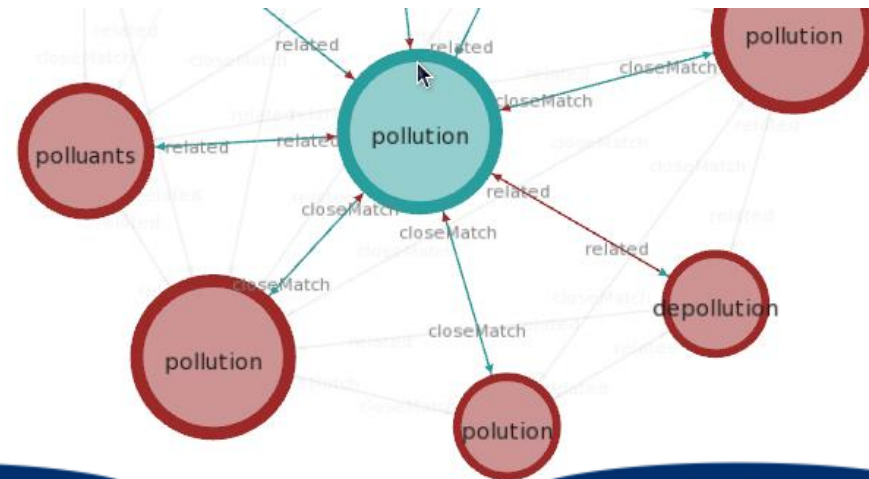
## thesaurus



# Combining metrics

## Orthographic metric

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



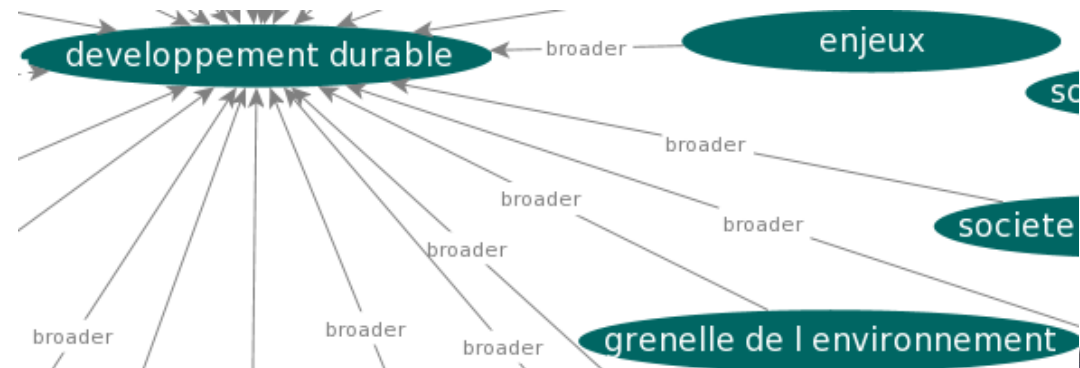
## Contextual metric

cosine vectors co-occurrent tags



## Social metric

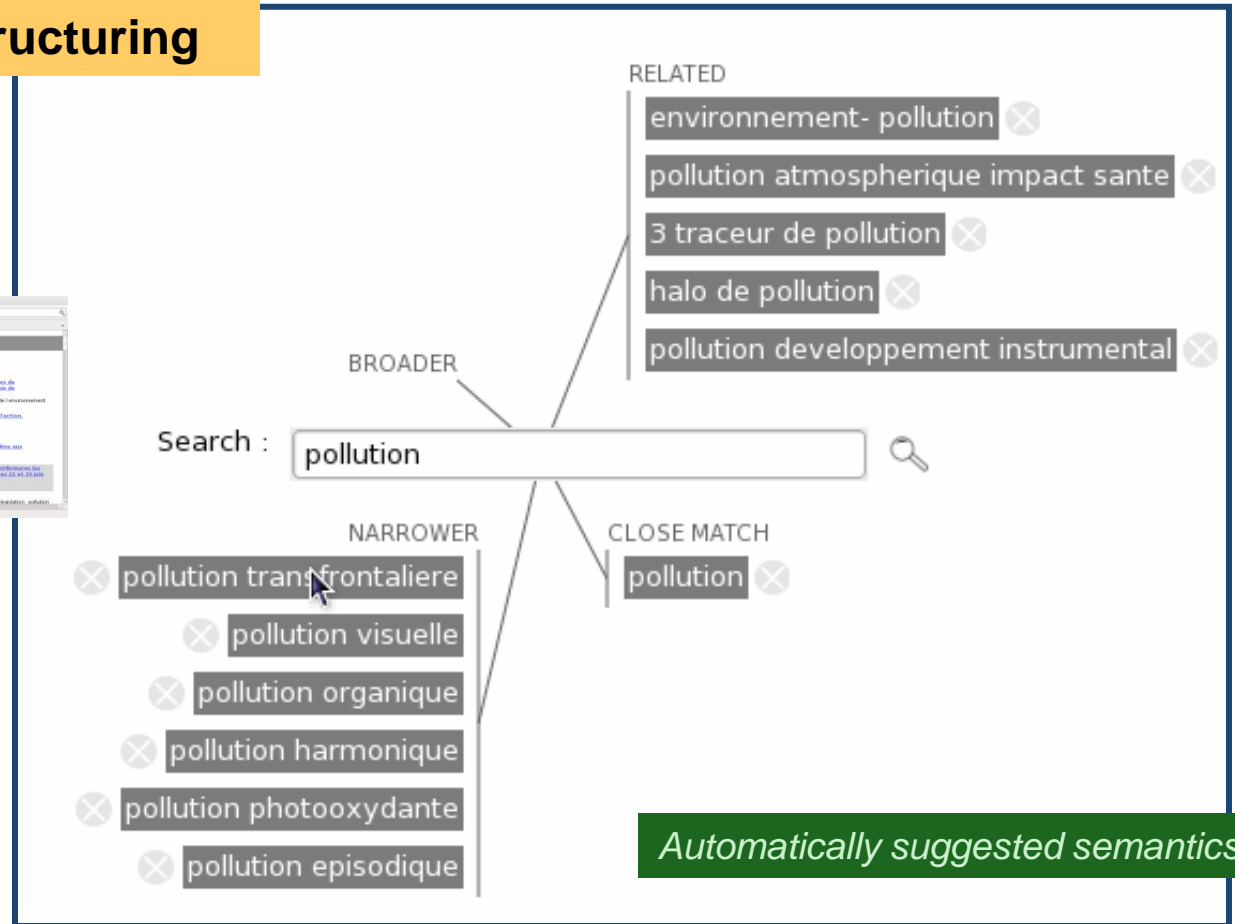
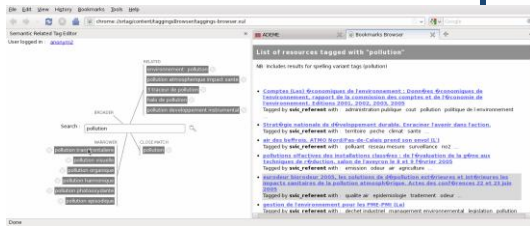
include user communities



# Interaction Design of SSW Applications: Tag Editor/Browser

## Involving users in tags structuring

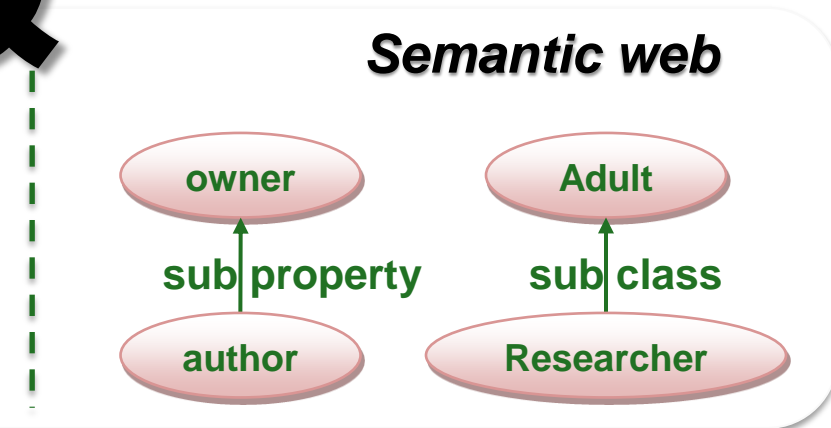
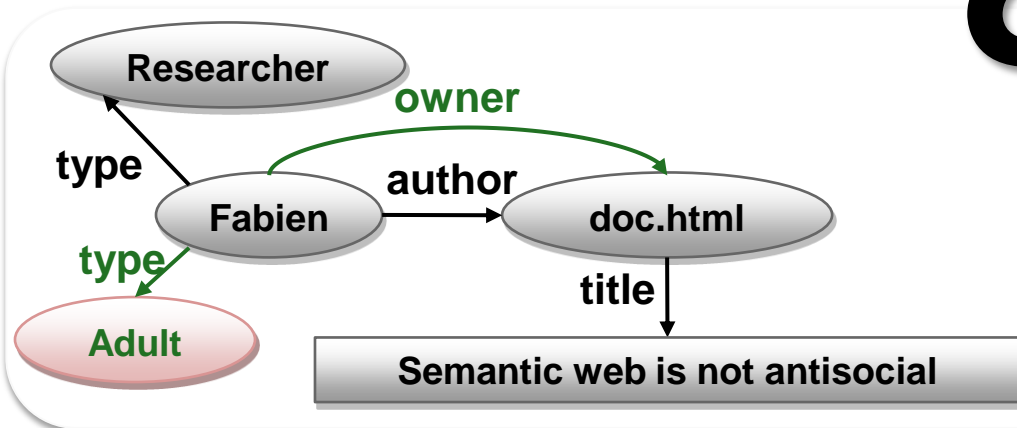
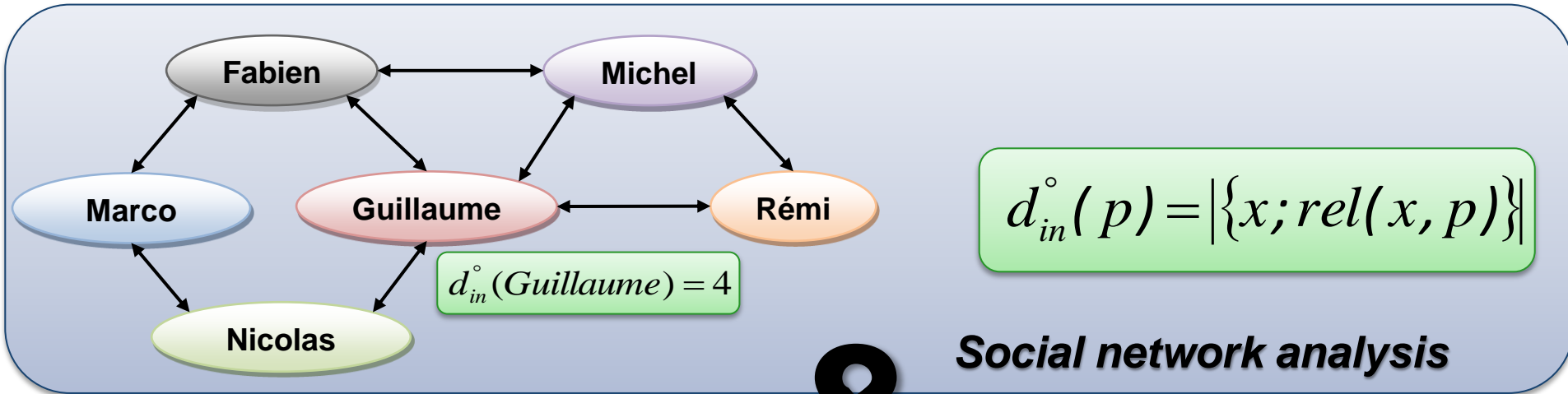
Context : ANR project **ISICIL**



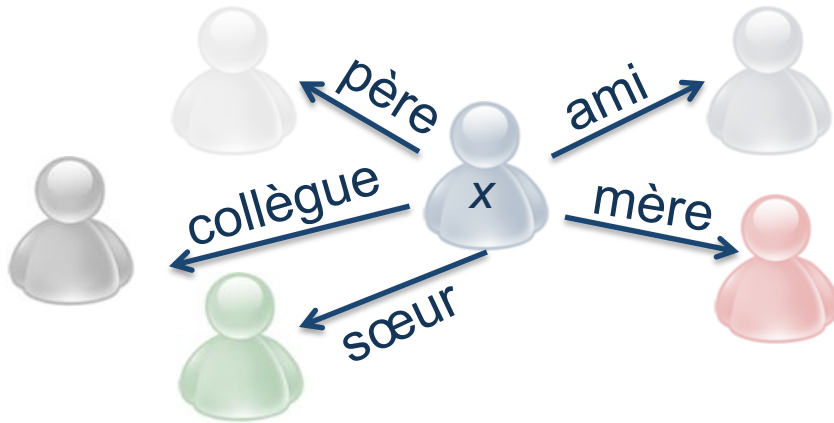
Automatically suggested semantics



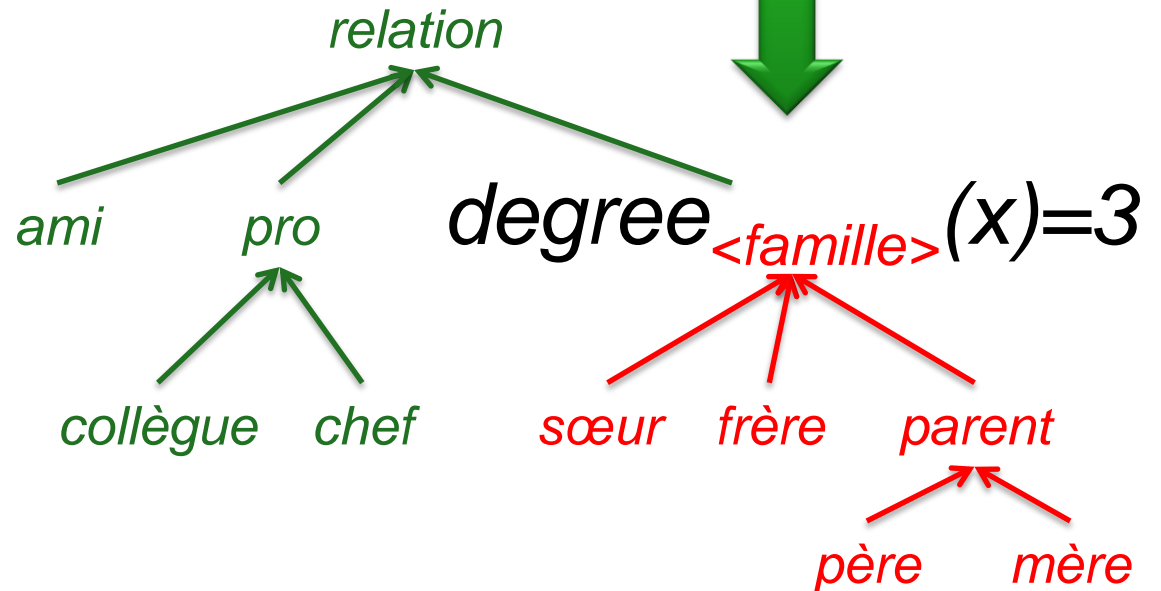
# Social Network Semantic Analysis



## ex. qualified degree



$degree(x)=5$



# Formal definition in SPARQL

SNA indices	SPARQL formal definition
$Comp_{\langle rel \rangle}(G)$	select ?x ?y from <G> where { ?x param[rel] ?y }group by any <sup>7</sup>
$D_{\langle rel, dist \rangle}(y)$	select ?y count(?x) as ?degree where { {?x (param[rel])*:: \$path ?y filter(pathLength(\$path) <= param[dist])} UNION {?y param[rel]:: \$path ?x filter(pathLength(\$path) <= param[dist])} }group by ?y
$D_{\langle rel, dist \rangle}^{in}(y)$	select ?y count(?x) as ?indegree where{ ?x (param[rel])*:: \$path ?y filter(pathLength(\$path) <sup>8</sup> <= param[dist]) }group by ?y
$Diam_{rel}(G)$	select pathLength(\$path) as ?length from <G> where { ?y s (param[rel])*:: \$path ?to }order by desc(?length) limit 1
$nb_{\langle rel \rangle}^g(from, to)$	select ?from ?to count(\$path) as ?count where{ ?from sa (param[rel])*:: \$path ?to }group by ?from ?to
$nb_{\langle rel \rangle}^g(b, from, to)$	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
$C_{\langle rel \rangle}^c(y)$	select distinct ?y ?to pathLength(\$path) as ?length (1/sum(?length)) as ?centrality where{ ?y s (param[rel])*:: \$path ?to     ... }group by ?y

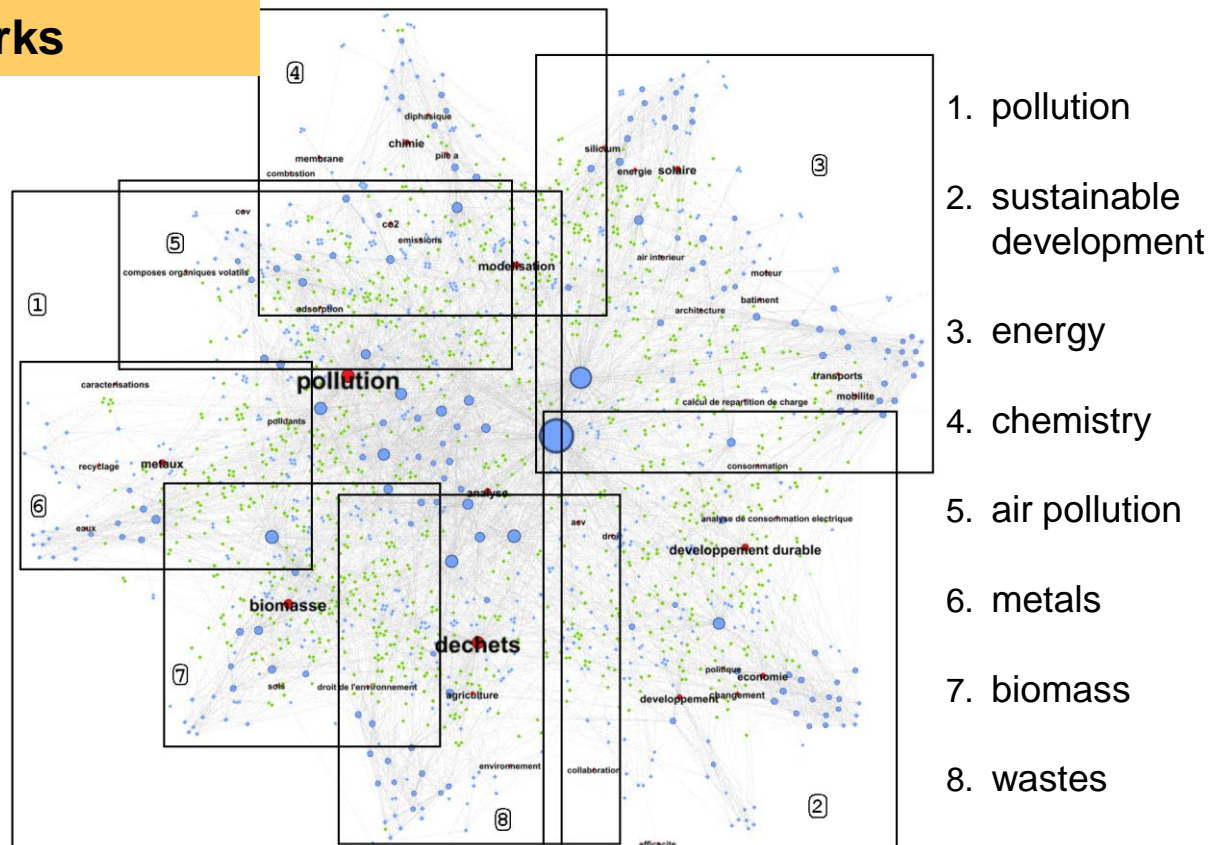


**CORESE/KGRAM**

# Interaction Design of SSW Applications: ISICIL Network Visualizer/Recognizer

## Recognizing and visualizing communities in social networks

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



*Mixed graph of actors and interests*

# Selected Publications

- Erétéo et al. Semantic Social Network Analysis, a Concrete Case, **Handbook** of Research on Methods and Techniques for Studying Virtual Communities, IGI Global, 2011.
- Gandon et al. Semantic Annotation and Retrieval: RDF, in **Handbook** of Semantic Web Technologies, Domingue, Fensel, Hendler, Springer; 1st Edition (June 24, 2011)
- Gandon, Ontologies in Computer Science, in Ontology Theory, **Book chapter** Management and Design: Advanced Tools and Models, IGI Global, 2010.
- Corby et al. The KGRAM Abstract Machine for Knowledge Graph Querying. **IEEE/WIC/ACM** 2010
- Erétéo et al. Analysis of a Real Online Social Network using Semantic Web Frameworks. **ISWC** 2009
- Buffa et al. SweetWiki: A Semantic Wiki. **Journal of Web Semantics**, 2008
- Khelif et al. An Ontology-based Approach to Support Text Mining and Information Retrieval in the Biological Domain. **Journal of Universal Computer Science**, 2007

## 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, Tadoo,  
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)

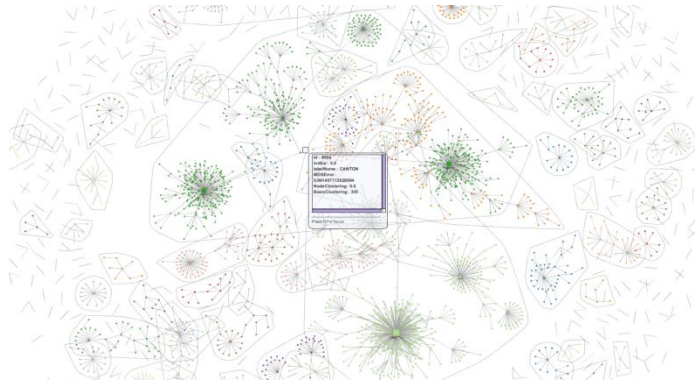
## Assistants

Christine Foggia (INRIA) and Marie H el ene Prosillico (I3S)

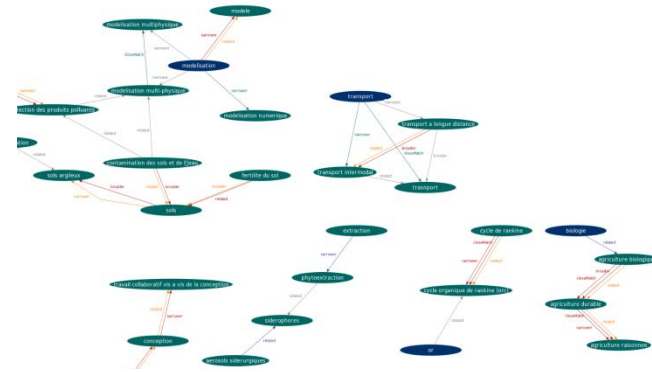
## PhD students

1. Pavel Arapov, 1st year (EDSTIC-INRIA)
2. Adrien Basse, 2nd year (UGB-INRIA)
3. Franck Berthelon, 2nd year (UNS-EDSTIC)
4. Ahlem Bouchahda, 3rd year (UNS-SupCom Tunis)
5. Khalil Riad Bouzidi, 2nd year (UNS-CSTB)
6. Luca Costabello, 1st year (INRIA-CORDI)
7. Corentin Follenfant, 1st year (SAP)
8. Maxime Lefran ois, 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)

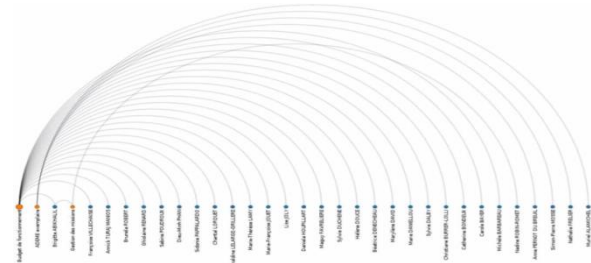
# graphs, graphs...



Ipernity social network structure



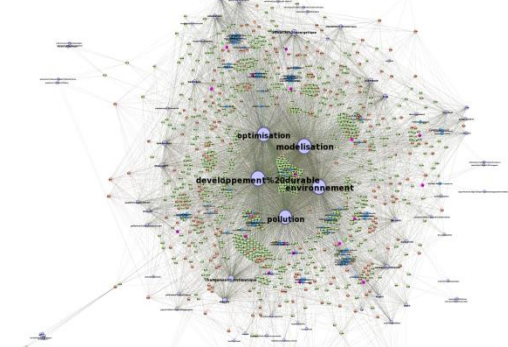
Folksonomy restructuring



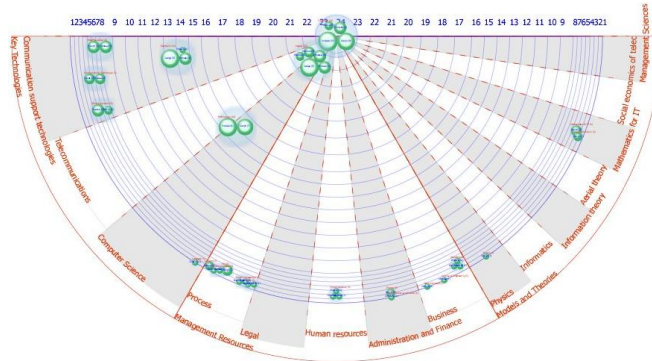
Interest graph



Interest matrix



Mixed graph of actors and interest



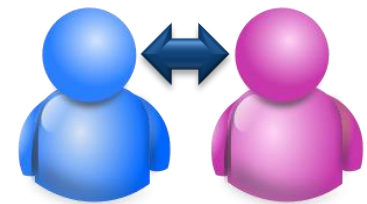
Clusters in KMP



# challenges

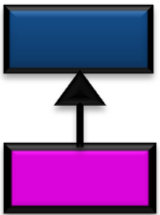
analyzing, modeling, formalizing and implementing graph-based social semantic web applications for communities.

multidisciplinary approach for analyzing and modeling



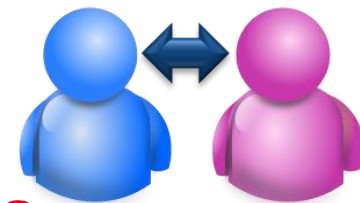
- the many aspects of intertwined information systems
- communities of users and their interactions

formalize and reason on these models



- new analysis tools and indicators
- new functionalities and better community management.

# interactions



- analyzing & modeling communities

interactions 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?

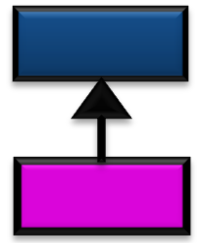
how do we reconcile and integrate the formalized stable semantics of computer science and the negotiable social interactions?

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)

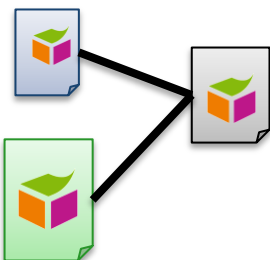
- enterprise social networking
- business intelligence, watching, monitoring
- communities of interest, of practice, of experts

## datalift.org (1/3)

- from 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

# the end

Thank you for your attention

