

Equipe SPARKS

Scalable and Pervasive softwARe and Knowledge Systems

ex GLC

Conseil du laboratoire I3S

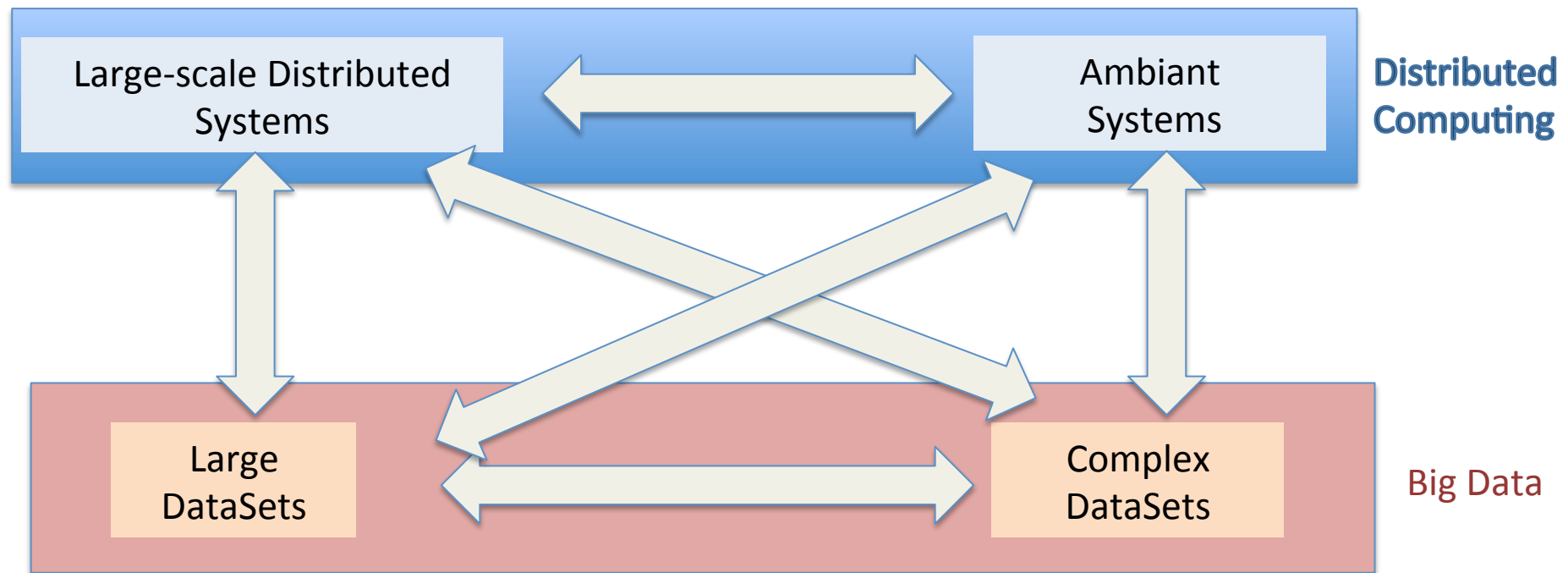
2/4/2015

- 1- Présentation générale & « administrative »
- 2- Présentation des 3 thèmes de l'équipe

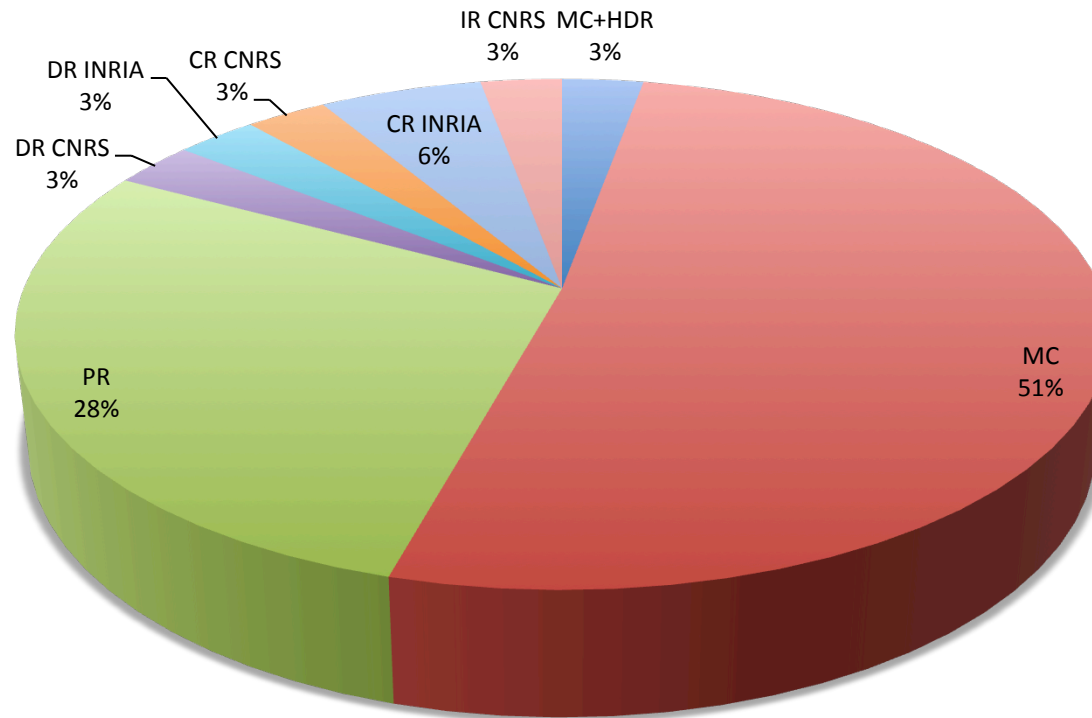
Big Picture : SPARKS

Two objectives :

- *“System-centric for preserving global viability” and*
- *“User-centric for incorporating local policies, adaptations, innovations”.*



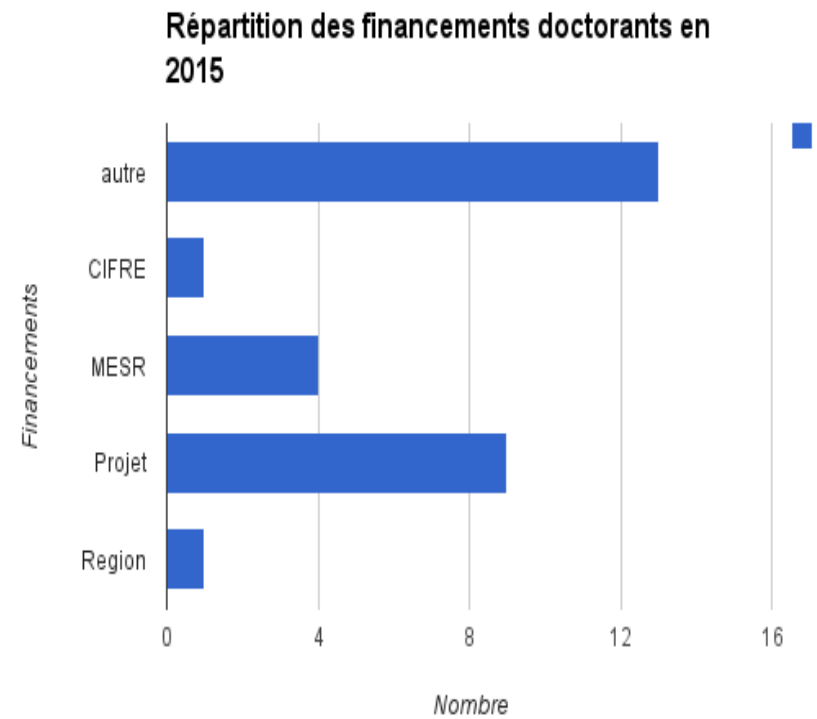
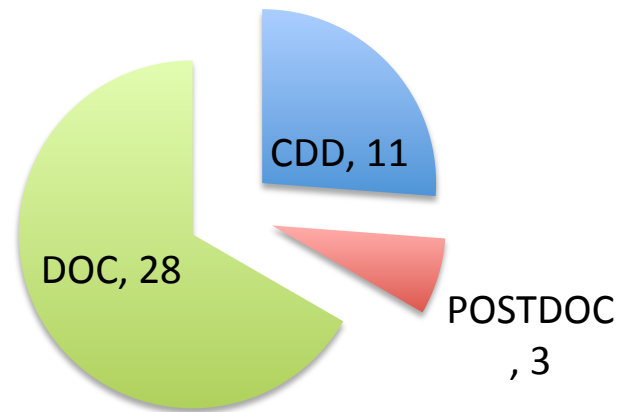
Membres permanents



2 PR associés.

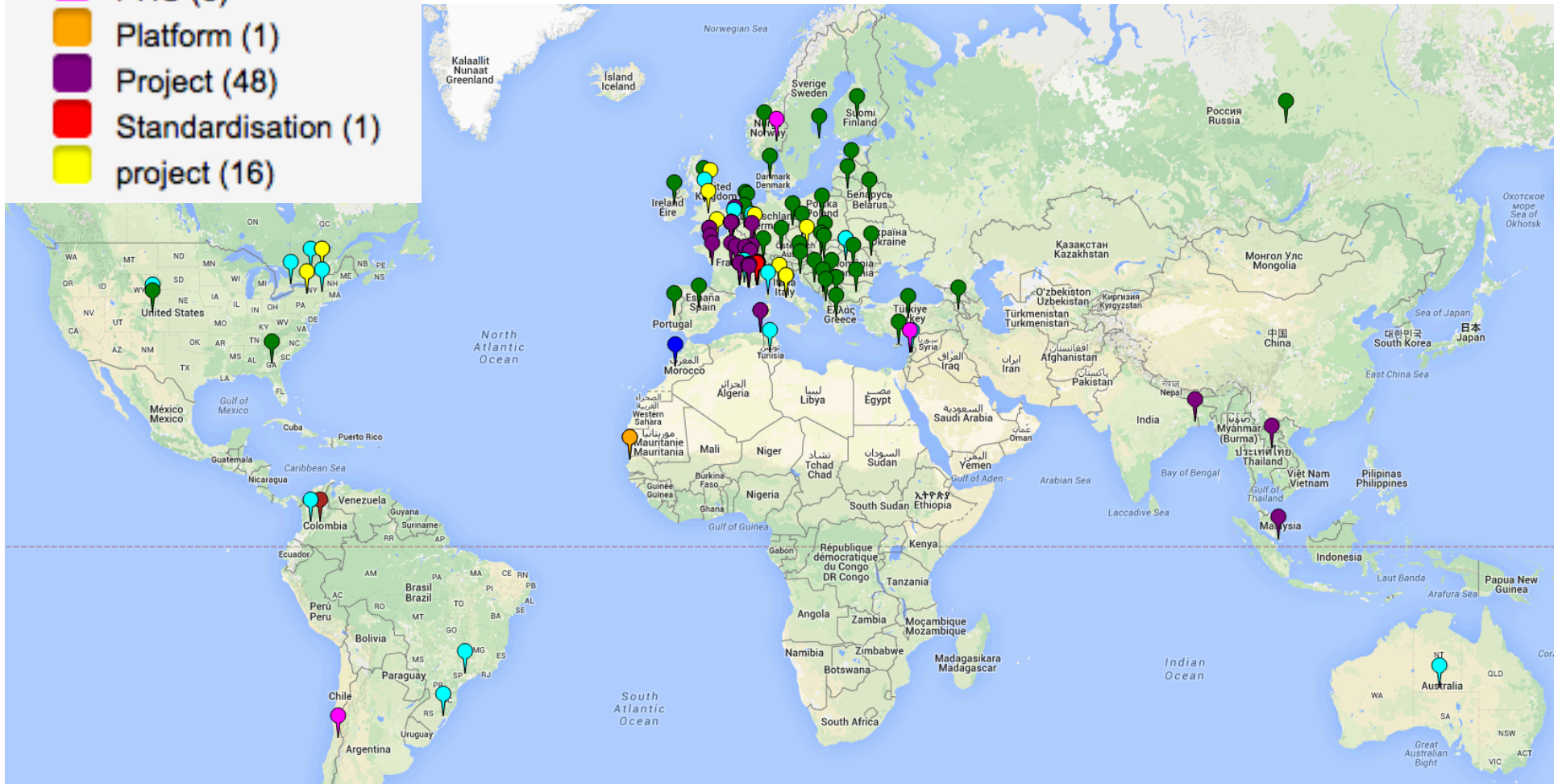
	MC+HDR	MC	PR	DR CNRS	DR INRIA	CR CNRS	CR INRIA	IR CNRS	Assistantes	Total
Total	1	18	8 (+2)	1	1	1	2	1	3	38

Membres non-permanents (2015)



- Co-direction (8)
- Co-tutelle (2)
- Collaboration (25)
- EU-Project (40)
- PHC (3)
- Platform (1)
- Project (48)
- Standardisation (1)
- project (16)

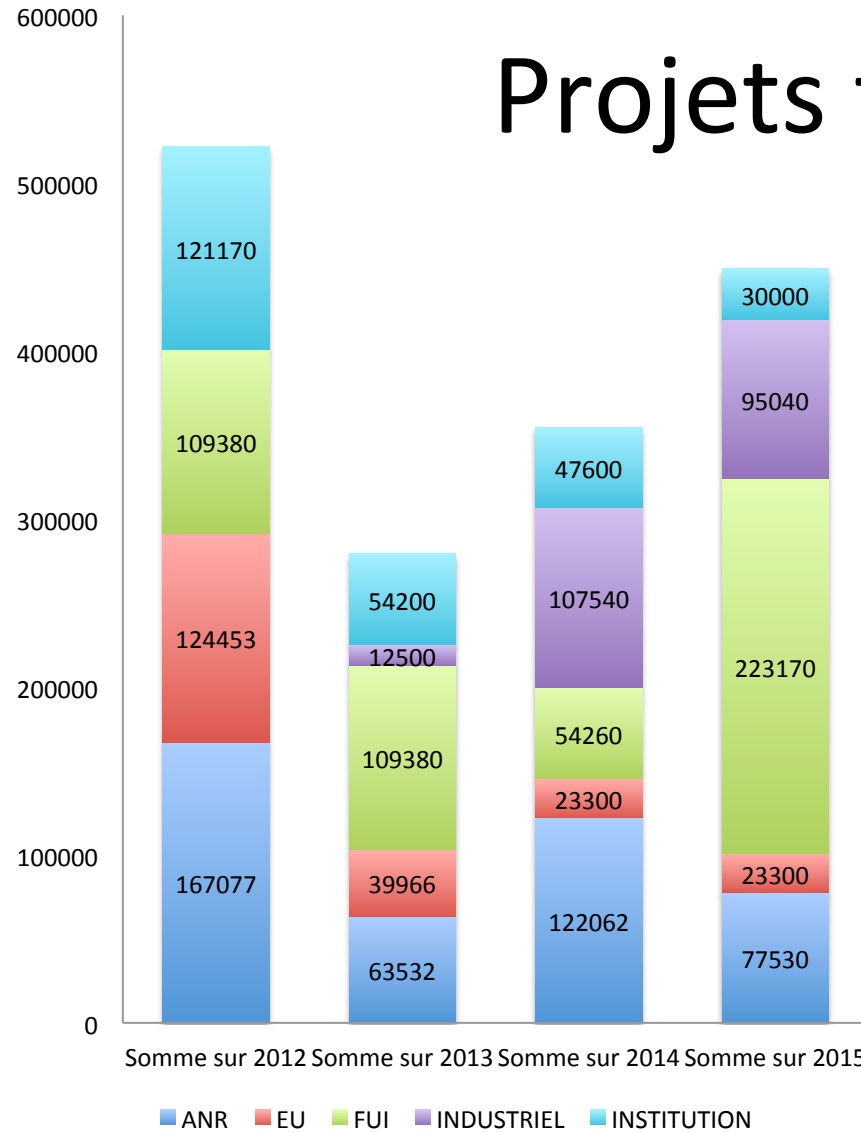
Collaborations



Collaborations inter-disciplines

	Académiques	Industrielles
Santé	CHU (groupe de travail : E-Santé de Proximité), CIU-Santé, INSERM, LPC	NEURELEC (implant cochléaire), Sagevox, ELA MEDICAL, Mnemotix, Gnubilla
Electronique	LEAT	ST-Microelectronic
Sciences du vivant	Institut de chimie de Nice, Conservatoire botanique national méditerranéen de Porquerolles, PNR des Préalpes d'Azur, Laos, ...	
SHS	MSHS (Axe 2 - TIC, usages et communautés, Axe 1 – Cognition et Coopération), BCL CPAM, GREDEG, LAPCOS	LudoTIC

Projets financés



CT	2015	TYPE
MASTODONS	30000	INSTITUTION
CONTINUUM	19000	ANR
VISIIR	58530	ANR
4TRAX - PACA	17000	FUI
PADDOC - PACA	147400	FUI
ER-FLOW	23300	EU
4TRAX - BPI	37260	FUI
PADDOC - BPI	21510	FUI
SYNCHRONEXT	7000	INDUSTRIEL
SEMANTIC	34900	INDUSTRIEL
AUTODESK	53140	INDUSTRIEL

Administration de la recherche & Animation Scientifique : *quelques éléments*

CNU 27 : 1 PR, 2MC ; **Co-CNRS** : 1 MCF

Consortium/GDR

- Pilotage :

- EGI-eu, EGI-engage, France Grille,

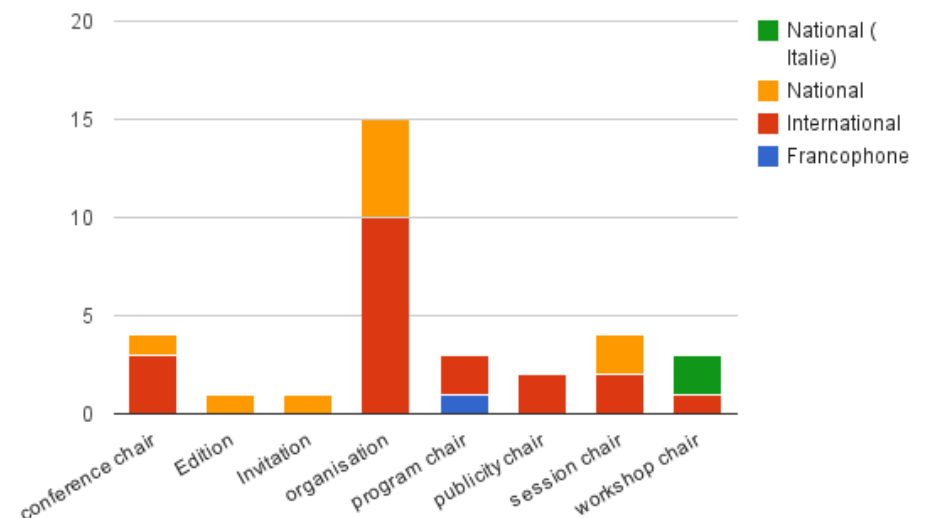
- GDR : GPL, MADICS,

- Actions: 1 GPL, 1 W3C, 2 ISIS

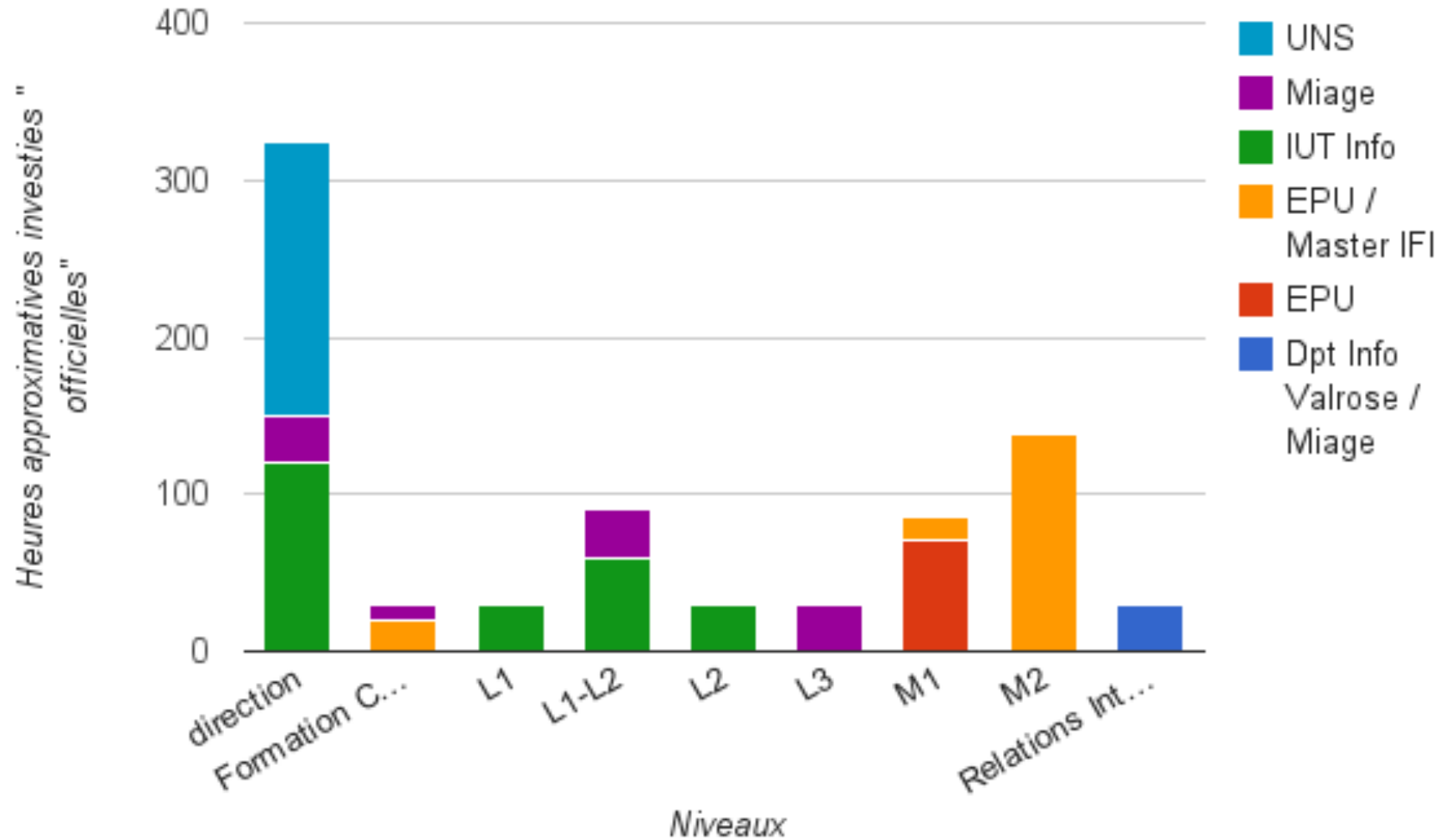
CA/CS Associations : CA/CS SIF, AFIA, Specif-Campus, Inforsid, ARCSIS, Pole SCS, PF
Conception, Incubateur Paca EST, ...

Responsabilités : Labex, labo,

Implication dans les Conférences



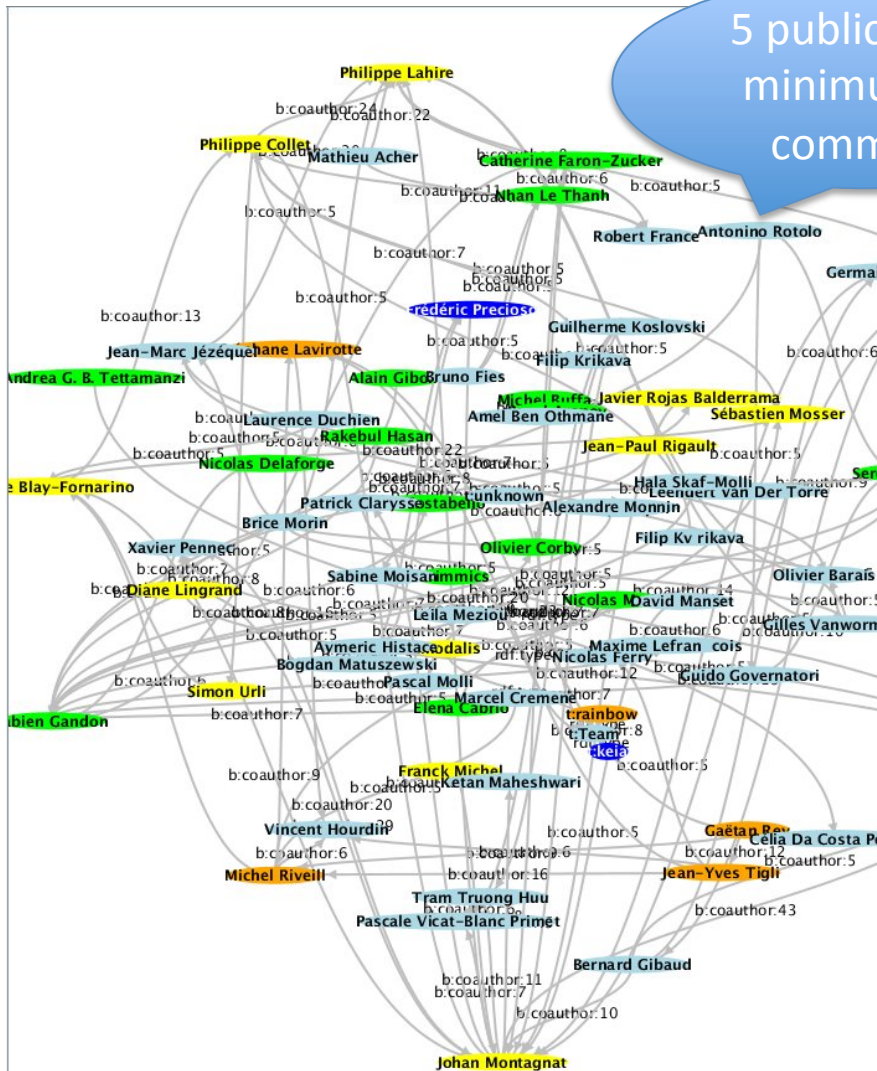
Implication des membres dans l'administration de l'enseignement 2014-2015



Publications

5 publications minimum en communs

Reconstruction du Réseau social de SPARKS par ses publications en utilisant CORESE



	TOTAL SPARKS					TOTAL
	2011	2012	2013	2014	2015	
rev-int	5	19	19	12	1	56
rev-nat	3	1	3	1	0	8
conf-int	57	71	46	39	2	215
conf-nat	7	23	2	6	0	38
conf-sa	3	30	0	0	0	33
inv-conf	0	0	0	0	0	0
livres-edit	0	3	4	4	1	12
livres-chap	0	7	11	15	1	34
logiciels	1	2	3	1	0	7
patents	0	0	0	0	0	0
these	0	4	4	9	0	17
HDR	0	1	0	0	0	1
Rapports	6	5	10	4	0	25

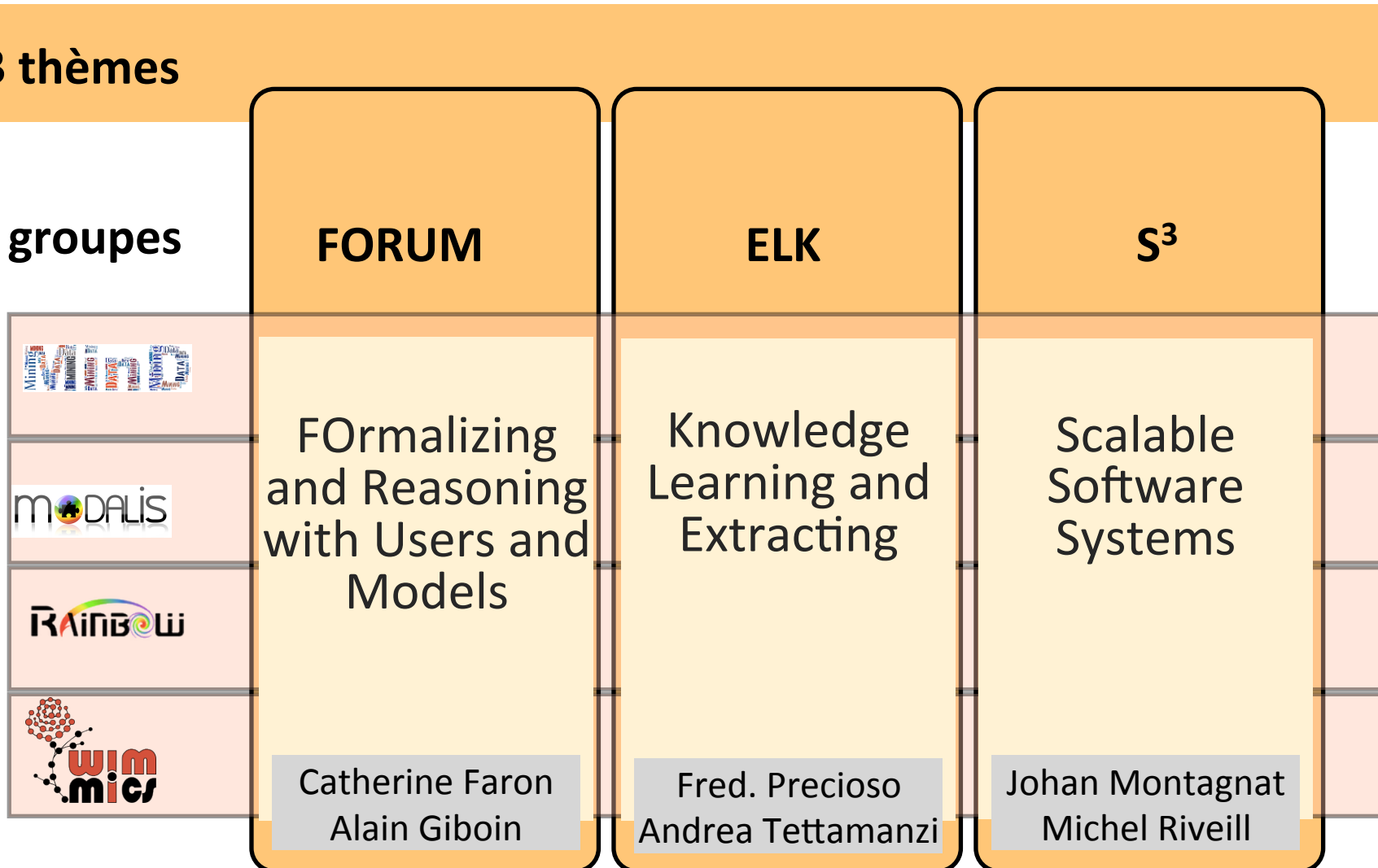
Vie de l' équipe

- Un comité des «représentants» d'équipe
 - ➔ réunions régulières : discussion sur le budget, sélection des projets, partage d'infos, ...
- 2 journées annuelles
- Séminaires croisés réguliers
- Fêtes
- ...




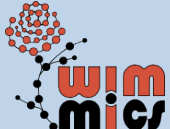
L'équipe SPARKS

3 thèmes

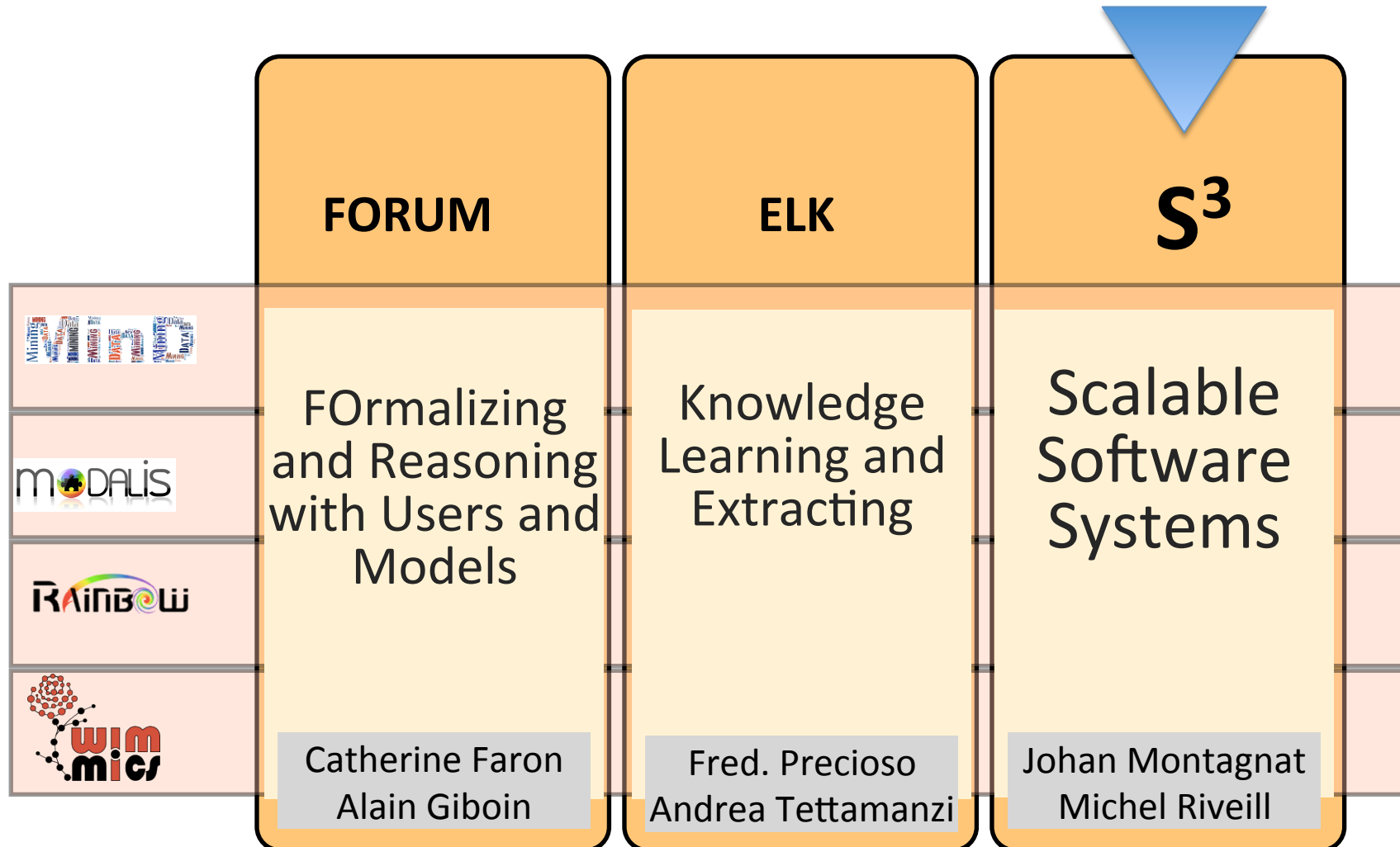
4 groupes



Des groupes aux thèmes

Contribution Utilisation	FORUM	ELK	S ³
	<ul style="list-style-type: none"> - Agents cognitifs et révision de croyances - Intégration de données 	<ul style="list-style-type: none"> - Fouille de données et apprentissage 	<ul style="list-style-type: none"> - Passage à l'échelle de l'apprentissage - Apprentissage embarqué
	<ul style="list-style-type: none"> - Intégration de données liées hétérogènes - Raisonnements distribués 	<p>Apprentissage pour les lignes de produits logiciels</p>	<ul style="list-style-type: none"> - Modèles de calculs distribués - Workflows scientifiques - Variabilité et évolutions de logiciels complexes.
	<ul style="list-style-type: none"> - Modélisation et raisonnements sur les interactions utilisateurs 	<p>Apprentissage pour l'adaptation dynamique des application à l'environnement physique et utilisateurs</p>	<ul style="list-style-type: none"> - Architecture pour l'adaptation dynamique des applications - Composition des données : des sources à la visualisation
	<ul style="list-style-type: none"> - Modèles utilisateurs et communauté sur le web - Sys. à base de connaissances - Représentation à base d'ontologies 	<ul style="list-style-type: none"> - Construction automatique de bases de connaissances et ontologies 	<p>Interrogation de données liées distribuées</p>

SPARKS : S3

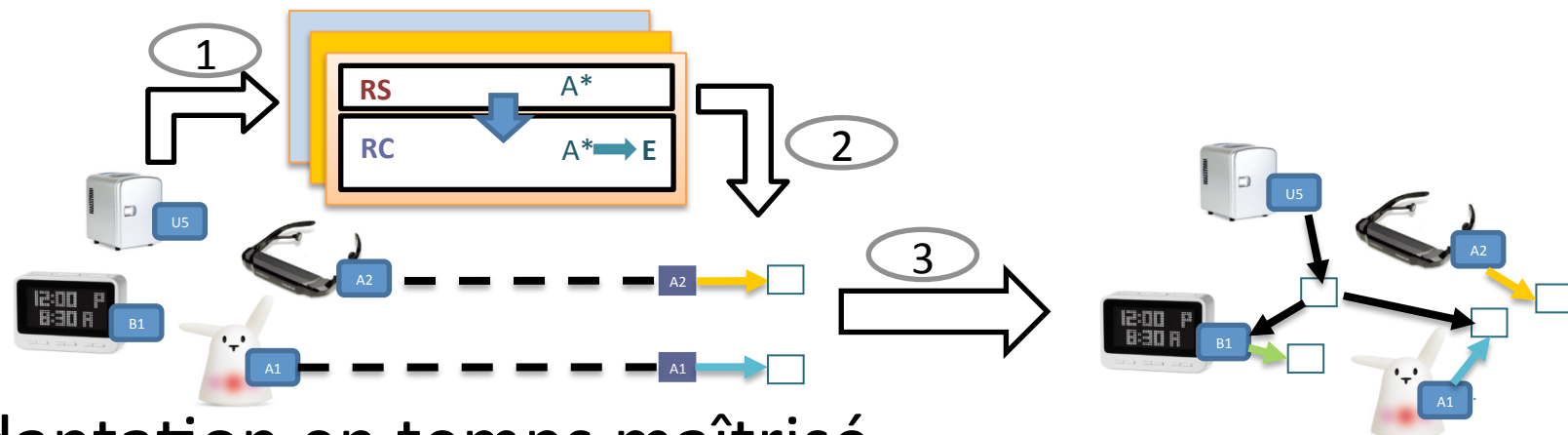


S³ – Scalable Software Systems

- **Distributed Systems, Ambient Systems:** omniprésence des systèmes distribués
 - Des fermes de calcul aux objets connectés
 - Données massives, liées et ouvertes disponibles
- **Défis principaux**
 - **Gestion de l'hétérogénéité**
 - Des systèmes (capacité variable des unités) → grilles, *clouds*, réseaux de capteurs
 - Des données (Sources de données distribuées) → FORUM
 - **Limiter les risques de l'interdépendance**
 - Propagation des pannes → auto-adaptativité, redondance
 - Adaptation à un environnement dynamique → recomposition

Adaptabilité

- Composition opportuniste fonction des services disponibles (*best effort*)
 1. Sélection des services prérequis par fragment
 2. Composition locale pour chaque fragment
 3. Composition globale



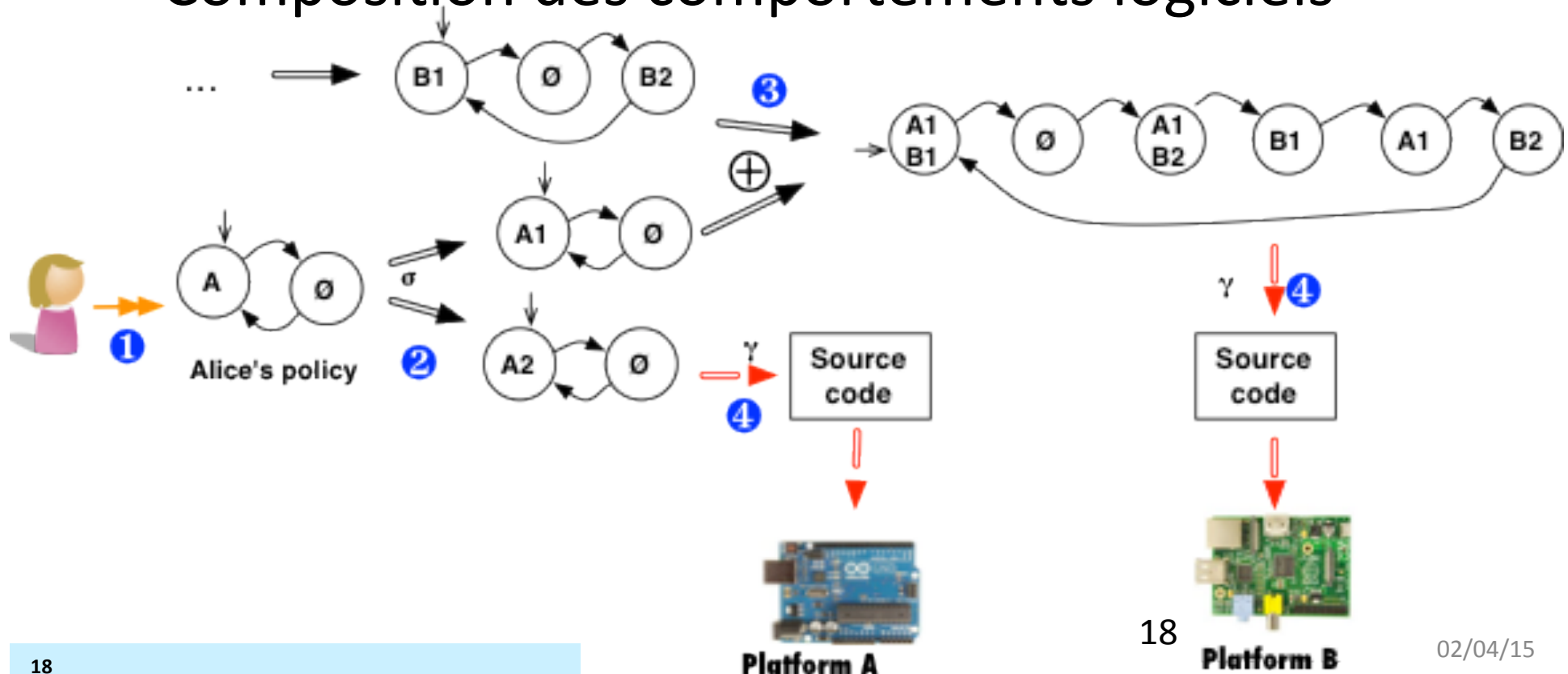
- Adaptation en temps maîtrisé
 - Pour maintenir la qualité de service

S³ – Scalable Software Systems

- Software (complexity)
 - Des réseaux de capteurs aux intergiciels
 - À faible couplage, assurant une forte disponibilité
 - Assemblages logiciels de grande taille
- Défis principaux:
 - **Gestion de la complexité logicielle** → composition logicielle, validité des assemblages logiciels *illustration*
 - **Gestion de la variabilité** → lignes de produits logiciels

Réseaux de capteurs

- Périphériques hétérogènes à capacité limitée
- Production continue de données à agréger
- Composition des comportements logiciels



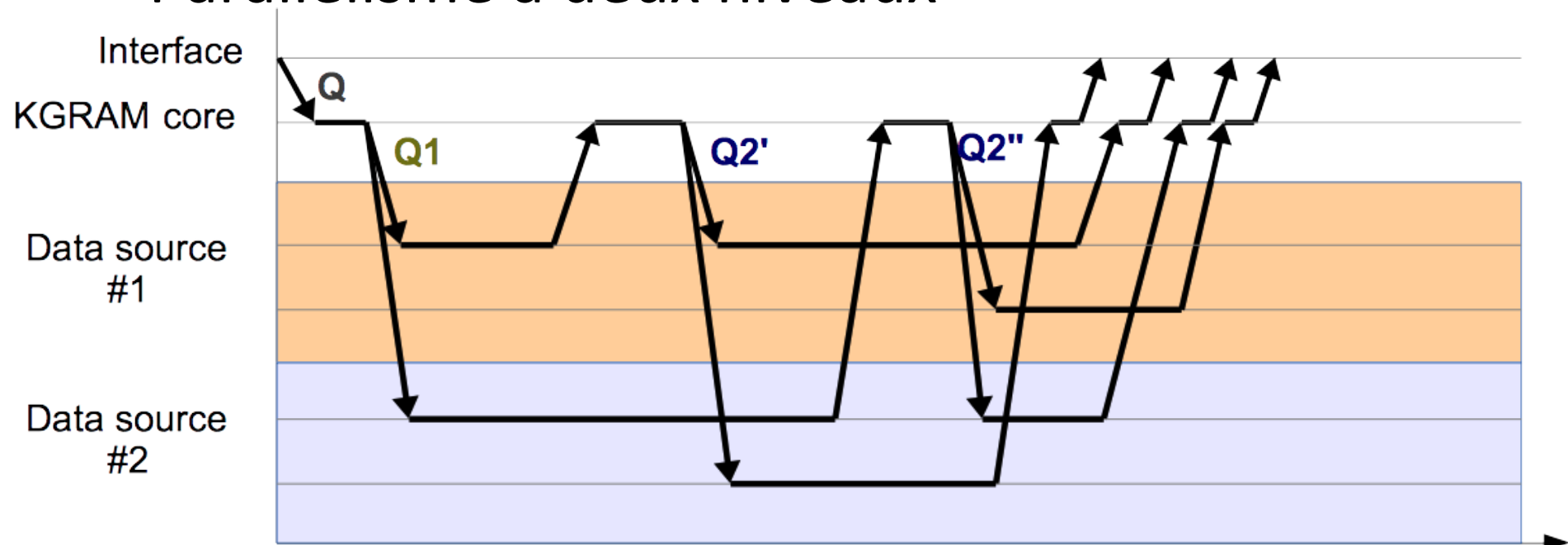
S³ – Scalable Software Systems

- **Scalability**, passage à l'échelle critique dans le cadre:
 - Expansion du nombre d'unités de calcul
 - Données massives distribuées
- Défis principaux
 - **Dégradation mesurée des performances** → modèles de performance, équilibrage de charge
 - **Traitement avec des moyens limités, logiciel embarqué** → prévision des temps de calcul, adaptation de la précision
 - **Flots de calculs dominés par les données** → langages de *workflows* scientifiques, flux de données
 - **Requêtes dans des données massives distribuées** → agrégation de flux de données, plans de requêtes distribués

illustration

Plans de requêtes distribués

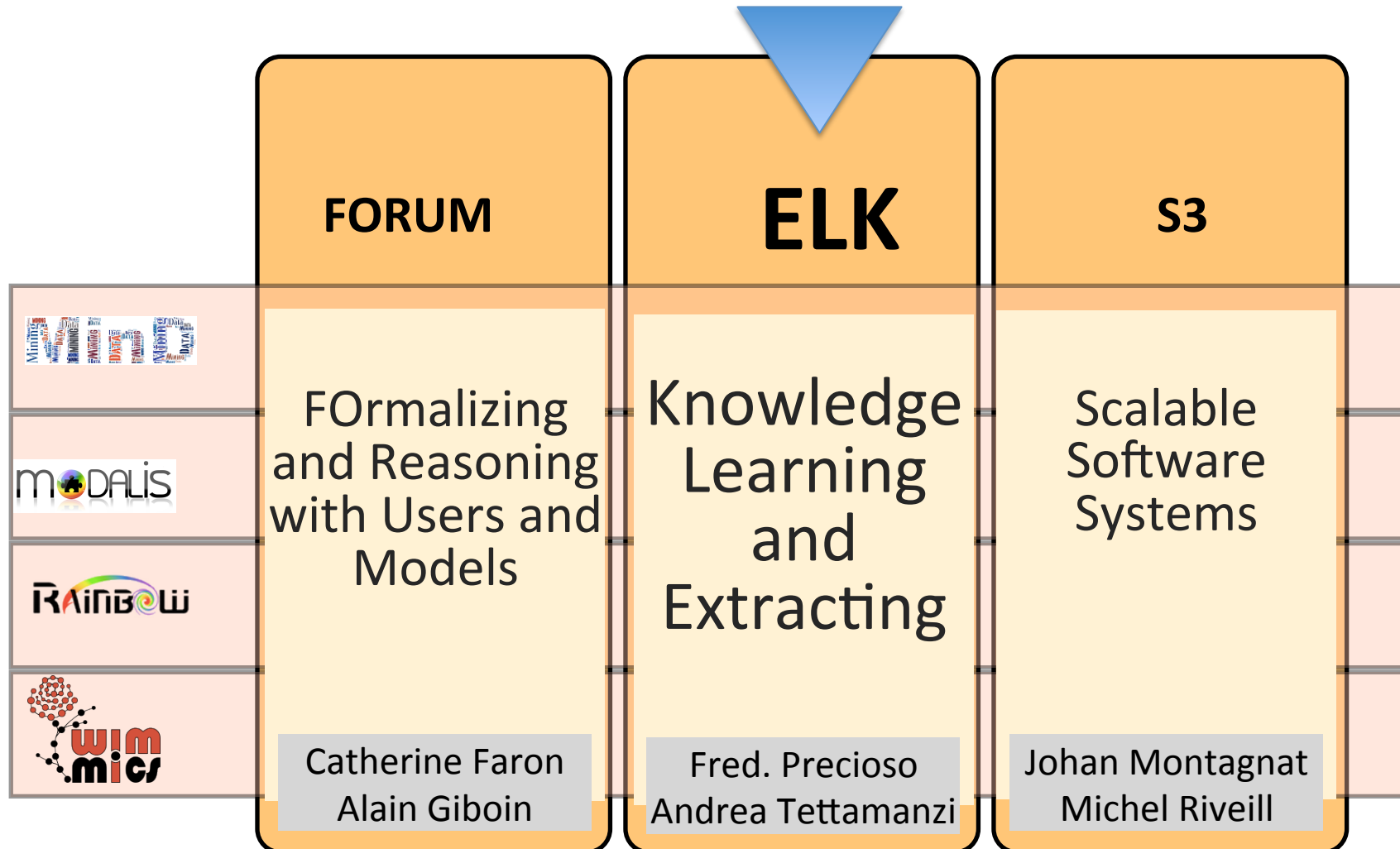
- Stratégies d'optimisation de requêtes
 - Recouvrement des communications, jointures distribuées
- Parallélisme à deux niveaux



Perspectives

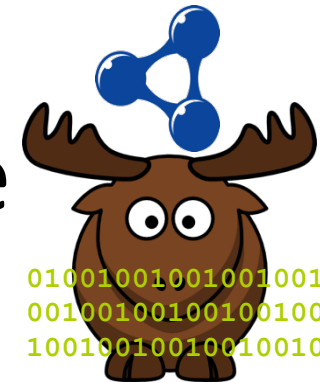
- Comportement « intelligent » émergent des compositions logicielles de grande taille
 - Véritable valeur ajoutée à la composition
- Passage à l'échelle des réseaux de capteurs
 - Campus → ville
- Virtualisation des plates-formes de calcul
 - Découplage logiciel / matériel
 - Reproductibilité des expériences, entrepôts de données « instrumentés » de moyens d'analyse
- Rapprochement des communautés « base de données » et « systèmes distribués »

SPARKS : ELK



ELK

Extracting and Learning Knowledge



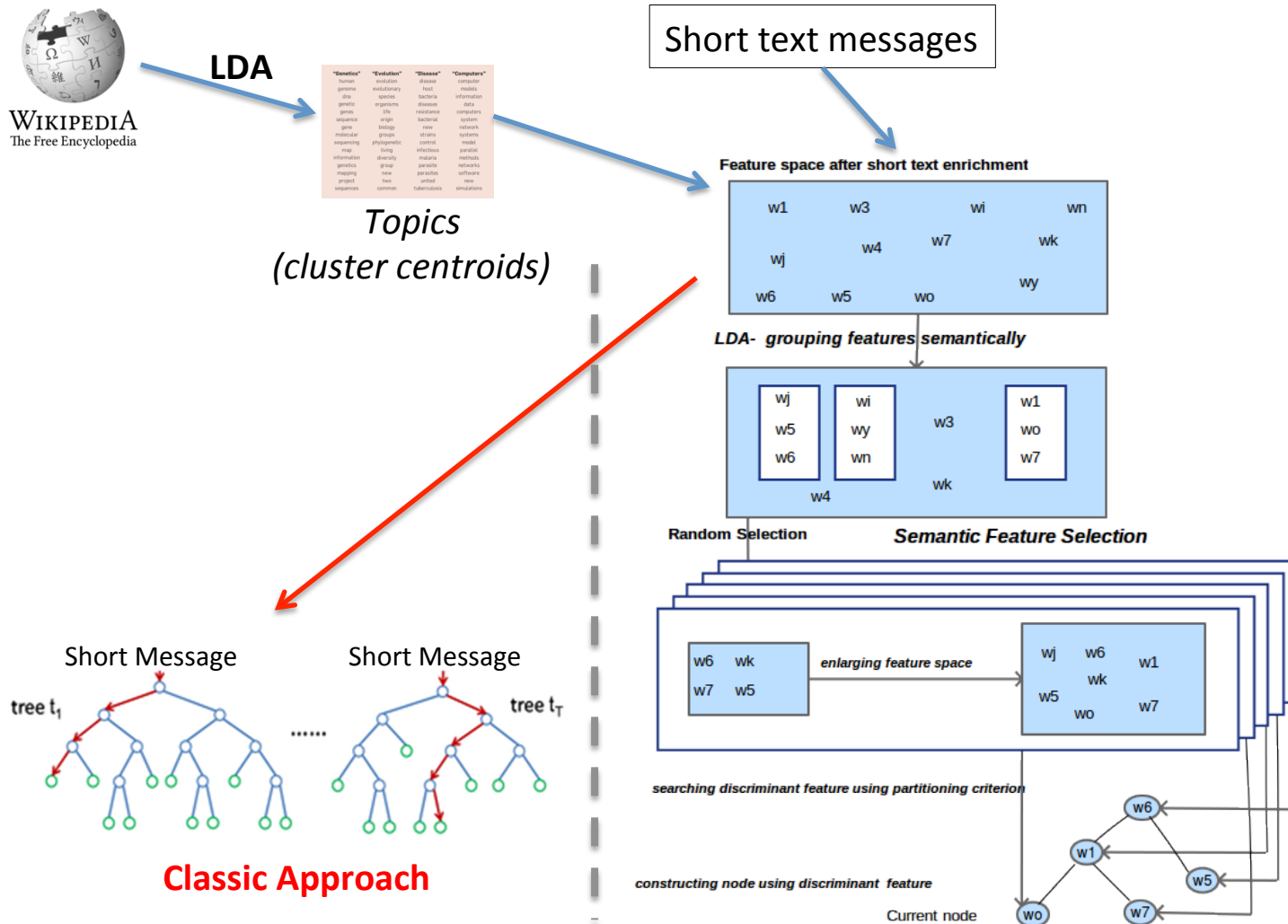
- Domaines
 - Machine Learning
 - Knowledge Discovery
 - Artificial Intelligence
- On conçoit des méthodes / algorithmes
- Enjeux
 - Passage à l'échelle (volume, embarqué, distribué)
 - Données hétérogènes (structurées / non-structurées)
 - Données \leftrightarrow Connaissances
 - Hybridation des méthodes
 - Traitement de l'incertitude et de l'imprécision

Fouille de textes

- Problématiques / Enjeux
 - La plupart des informations sont en format textuel
 - Extraire info & connaissances
 - Textes courts (SMS, *tweets*, etc.) : pas de contexte, peu d'informations
- Classification temps réel
- Classification de flux de textes
- Recherche interactive
- Travaux
 - Regroupement / classification de textes courts
 - Question answering (Qakis)
 - Zoomathia
 - Fouille d'opinions / sentiments

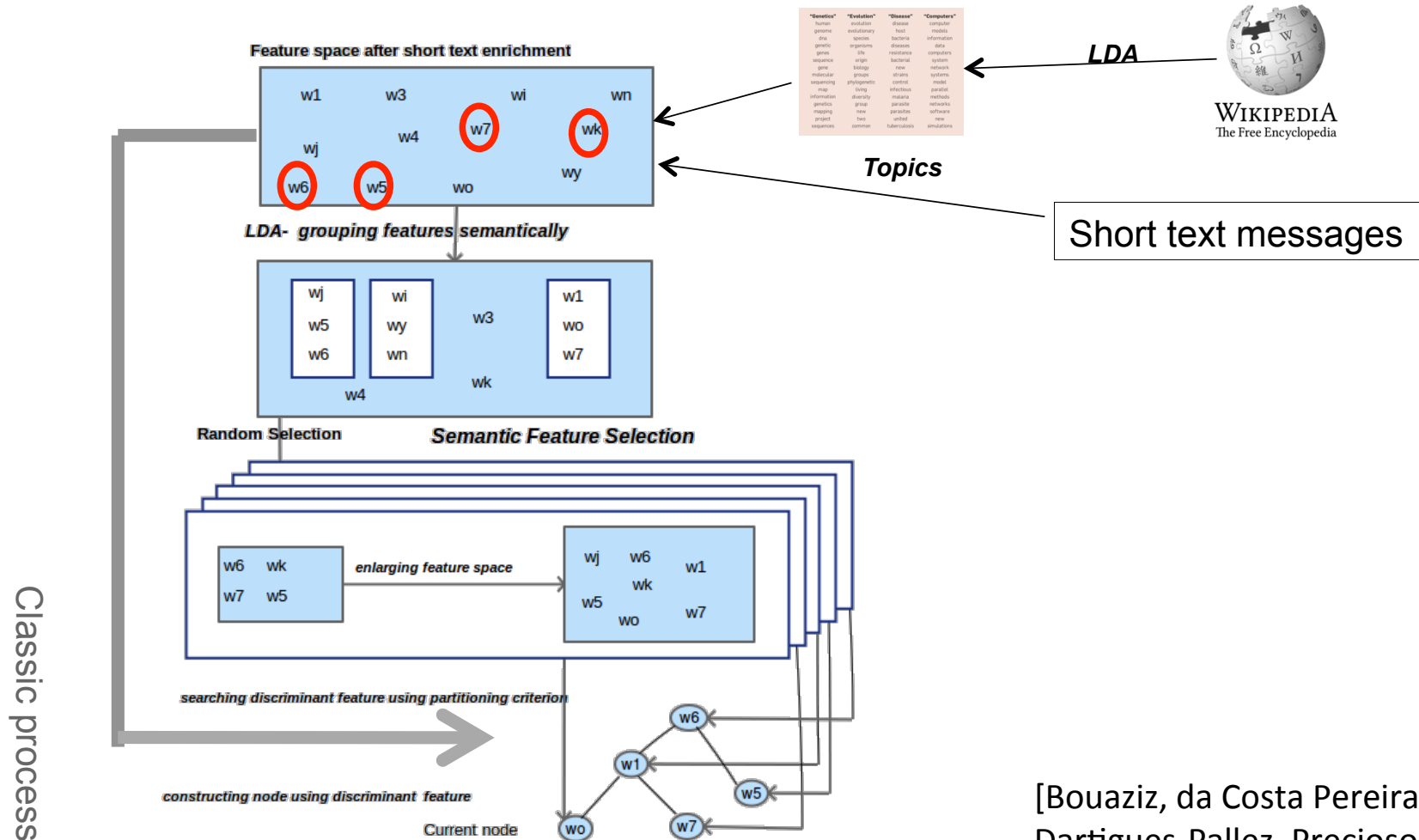
Fouille de textes

- Contribution : *Semantic Random Forest*



Fouille de textes

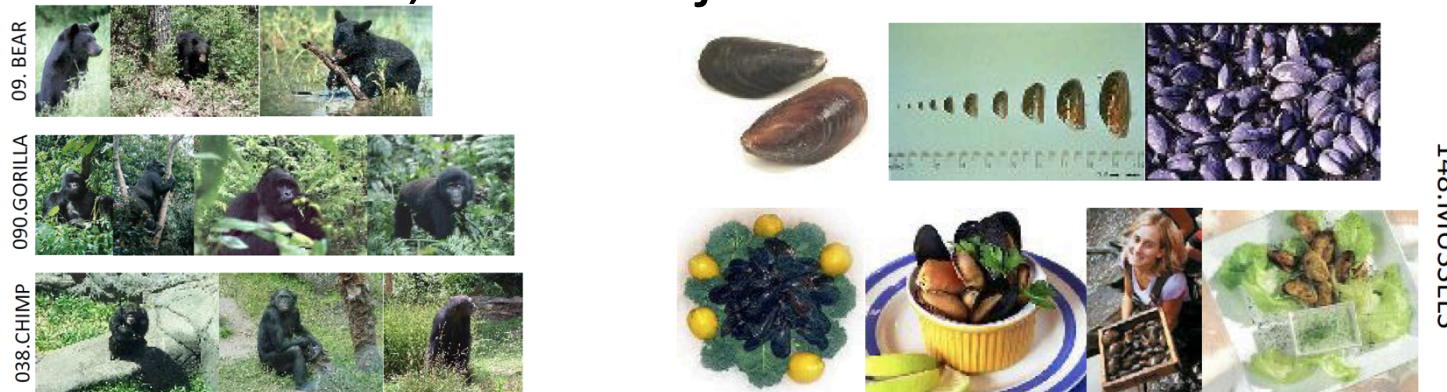
- Solutions : *Semantic Random Forest*



[Bouaziz, da Costa Pereira, Dartigues-Pallez, Precioso]

Fouille de données Multimédia

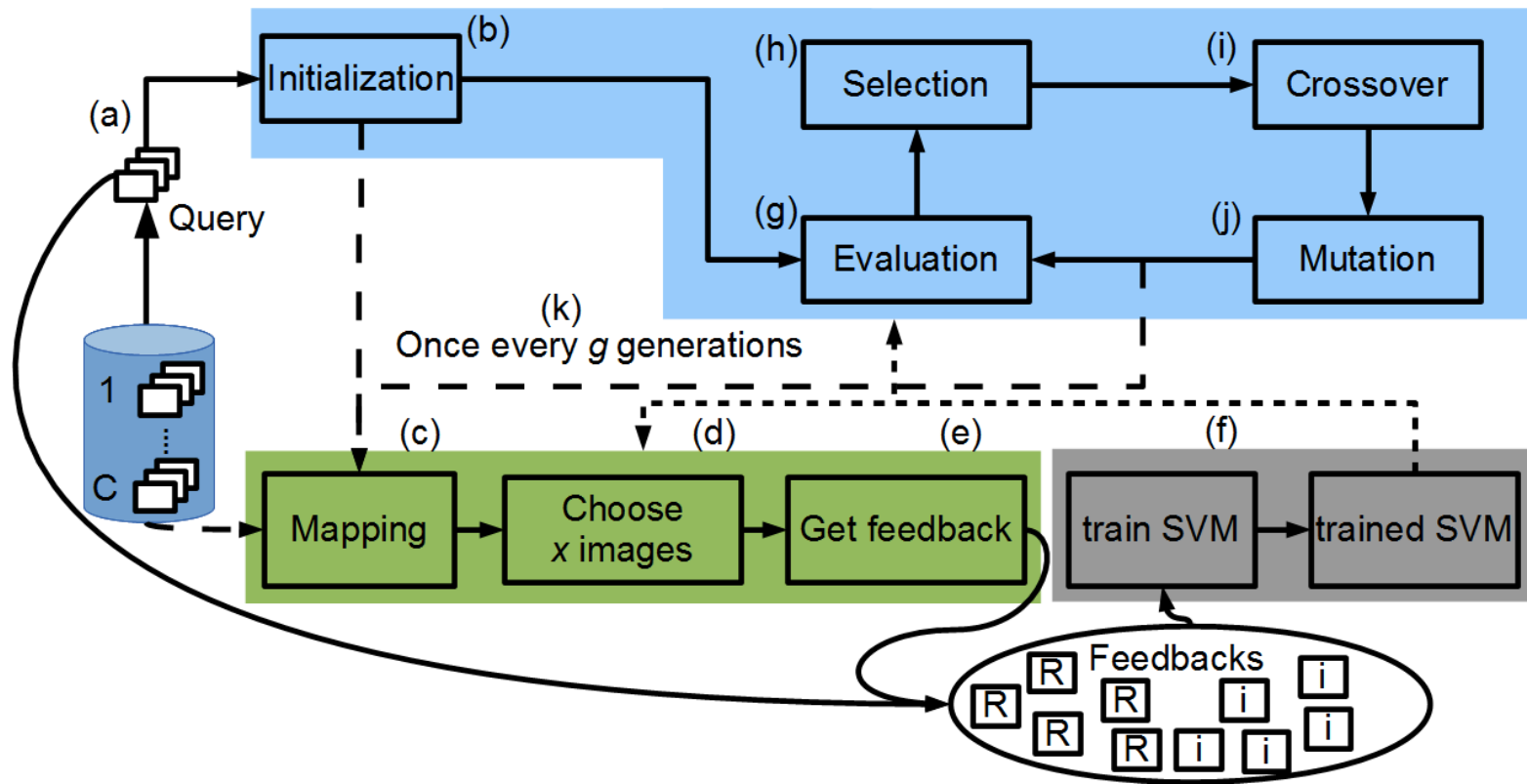
- Passage à l'échelle & Recherche interactive
- Multi-modalité, Multi-objectifs



- Apprentissage sur des objets complexes :
 - Dualité : dimension temporelle / dimension spatiale (ex. contenu vidéo)
 - combinaison de données structurées et non-structurées (image, son, vidéo + texte, GPS... + métadonnées)

Fouille de données Multimédia

- Solution : Hybridation (SVM, MOGA, LSH, Active Learning)

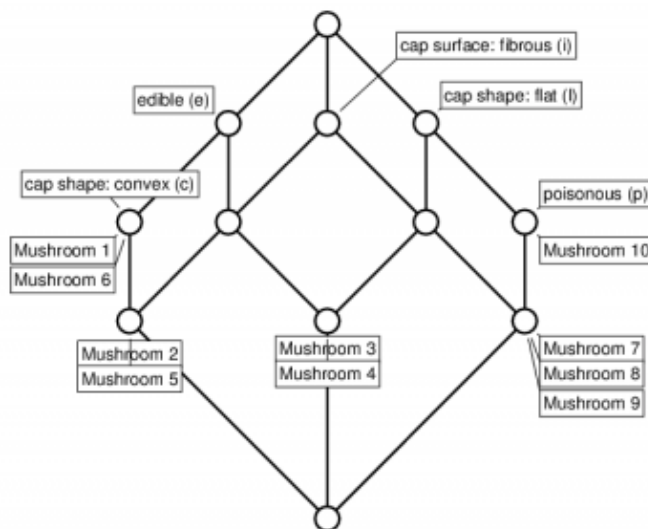


Knowledge Discovery

- Problématiques
 - Extraction d'informations et connaissances de données structurées
 - Entrepôts de données, *big data mining/analytics*
 - Web de données
 - Intégration de métadonnées et ontologies
- Directions
 - *Conceptual Biclustering*
 - Construction automatique d'ontologies

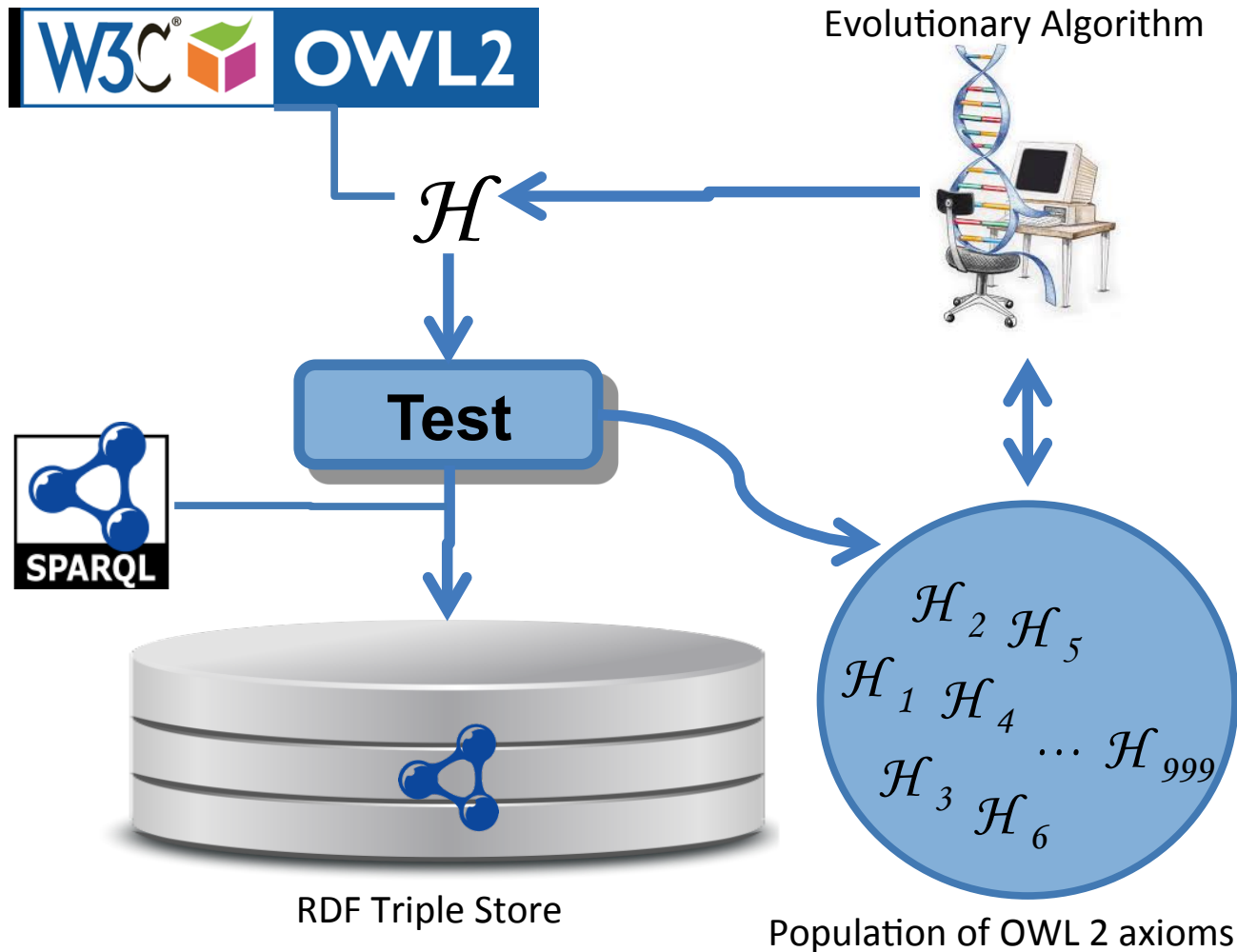
Conceptual Biclustering

	edible (e)	poisonous (p)	cap shape: convex (c)	cap shape: flat (f)	cap surface: fibrous (i)
Mushroom 1	X				
Mushroom 2	X		X		
Mushroom 3	X			X	
Mushroom 4	X				X
Mushroom 5	X		X		X
Mushroom 6	X			X	
Mushroom 7		X	X		X
Mushroom 8		X	X	X	X
Mushroom 9		X	X	X	X
Mushroom 10		X	X	X	X



- Treillis de concepts
- Cadre théorique pour
 - l'extraction de règles d'association
 - La génération de règles de classification
 - La classification conceptuelle double

RDF Mining



Research question:
How to automatically extract (OWL 2) ontological axioms from a set of known facts (an RDF dataset)?

Epistemological Problem:

- Test of hypotheses → falsification (cf. Popper)
- Noise in the data → possibilistic treatment

Search Problem:

- Evolutionary algorithms
- Relational Learning/ILP

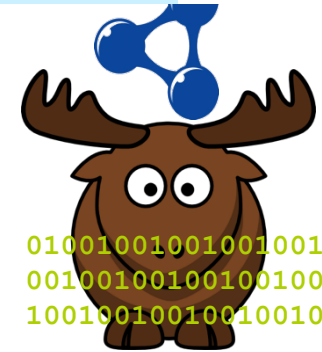
[Tettamanzi, Faron-Zucker, Gandon]

Ensemble Learning

- Apprendre séquentiellement ou simultanément des décisions partielles de l'environnement (ex. capteurs) et la décision globale
- Temps réel & Apprentissage interactif
- Apprentissage hiérarchique
- Apprentissage auto-adaptatif à l'évolution de l'environnement

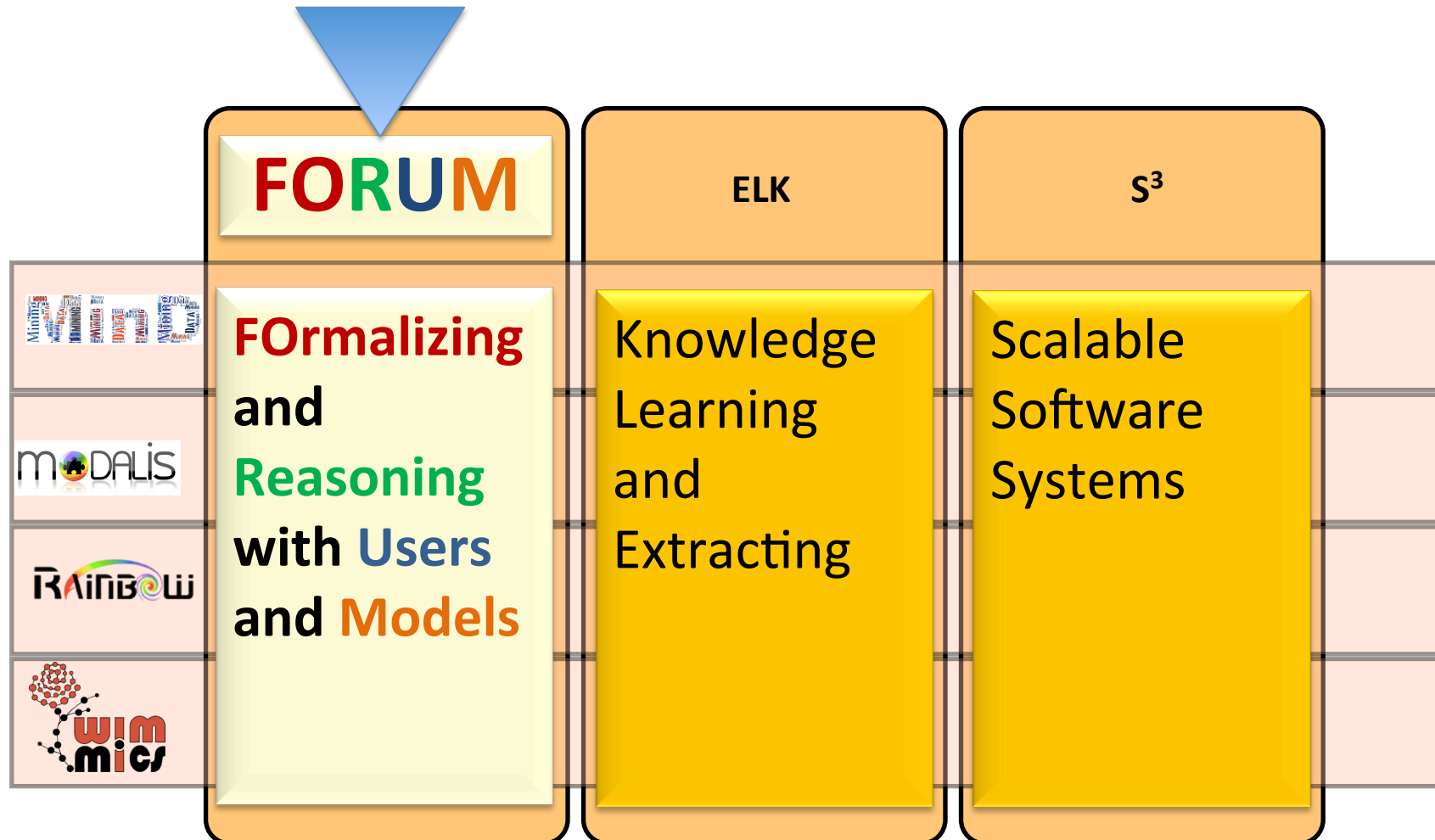
[Dartigues-Pallez, Precioso, Lavirotte, Tigli]

ELK - Perspectives



- Apprentissage interactif (mettre l'utilisateur au centre)
- Web sémantique (formalismes et représentation)
- Pluridisciplinarité (santé, linguistique, économie, géographie, ...)
- Transfert et valorisation

SPARKS : Forum



FORUM

F**o**r**m**alizing and R**e**asoning with U**s**ers and M**o**dels

Systèmes sémantiques/à base de connaissances
Systèmes interactifs

ENJEUX

Représentation et **raisonnements** à base d'**ontologies**

Données liées, hétérogènes

Raisonnements à base de graphes, distribués

Intégration de données

Interactions avec et entre les utilisateurs

Modèles d'utilisateurs, d'interactions, de communautés

Agents cognitifs et révision de croyances

Web collaboratif

FORUM

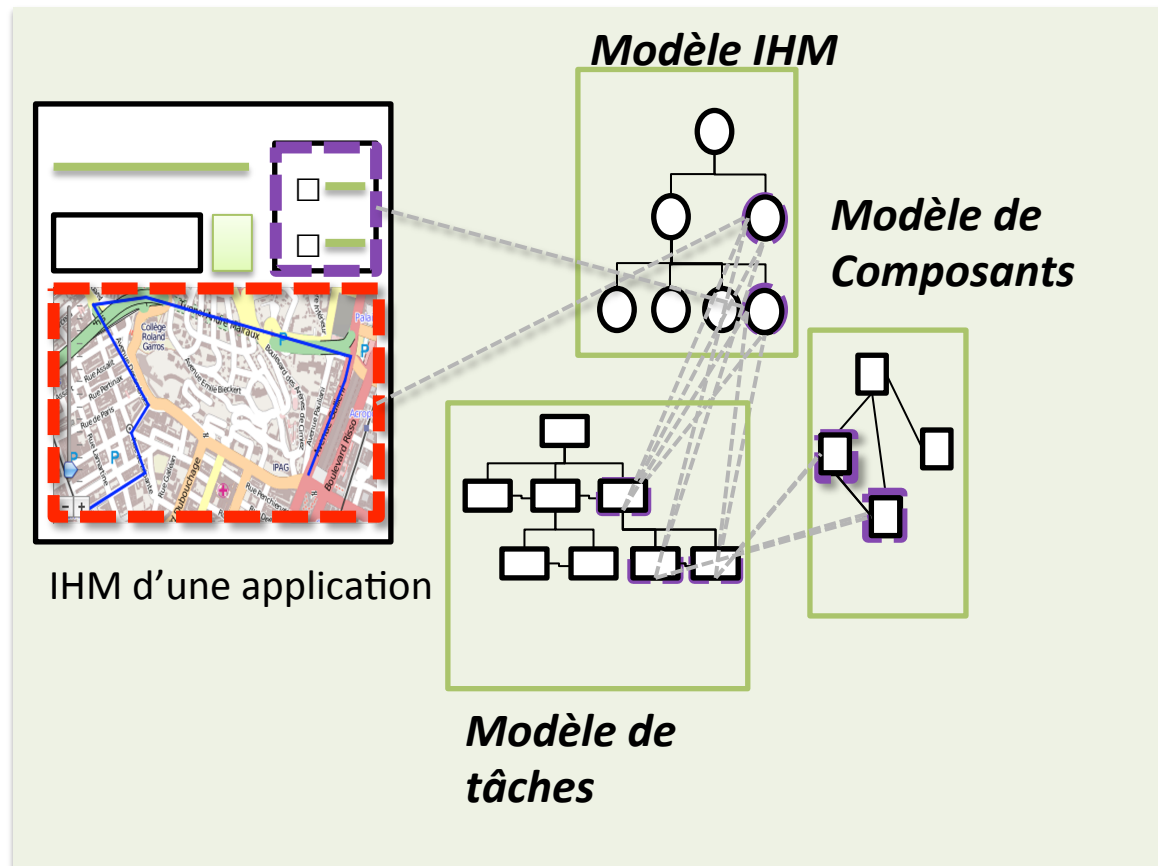
FOrmalizing and Reasoning with Users and Models

interactions avec les utilisateurs

OntoCompo:

Composition d'applications à base d'ontologies

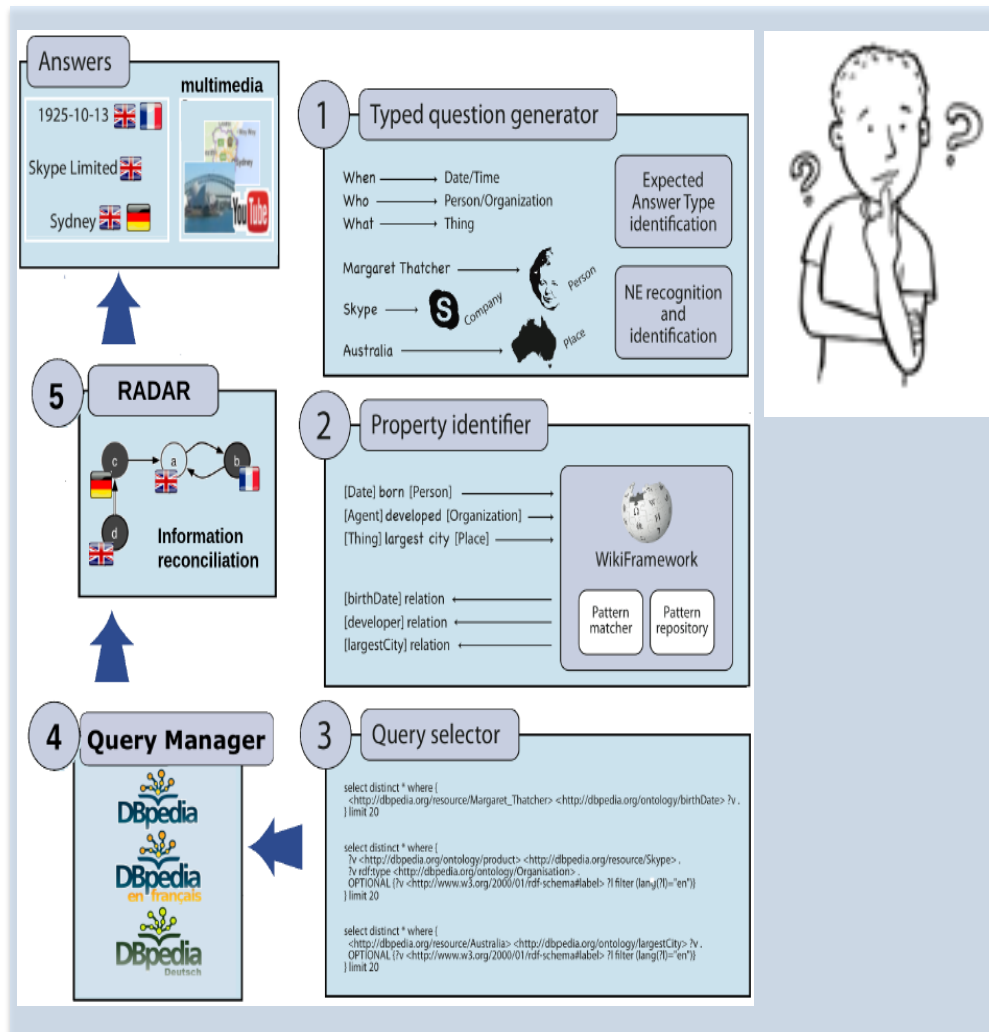
[Brel, Renevier-Gonin,
Giboin, Déry, Riveill]



- **Créer une nouvelle application opérationnelle** à partir d'applications existantes
- **Gérer l'interaction** <Fonctionnel x IHM x Usage>

- **Modélisation à base d'ontologies**
- **Extraction et assemblage à base d'ontologies**
- **Expérimentations** avec utilisateurs-développeurs (validation)

Systeme de Questions-Réponses [Cabrio, Cojan, Gandon]



- Répondre à des **questions** posées en **langue naturelle** par l'utilisateur

- Exploitation d'une BD structurée : **Dbpedia** (Wikipedia sémantique)
- Interprétation des questions basée sur le **matching des relations**
- Module d'argumentation **RADAR** (explication et justification des réponses)

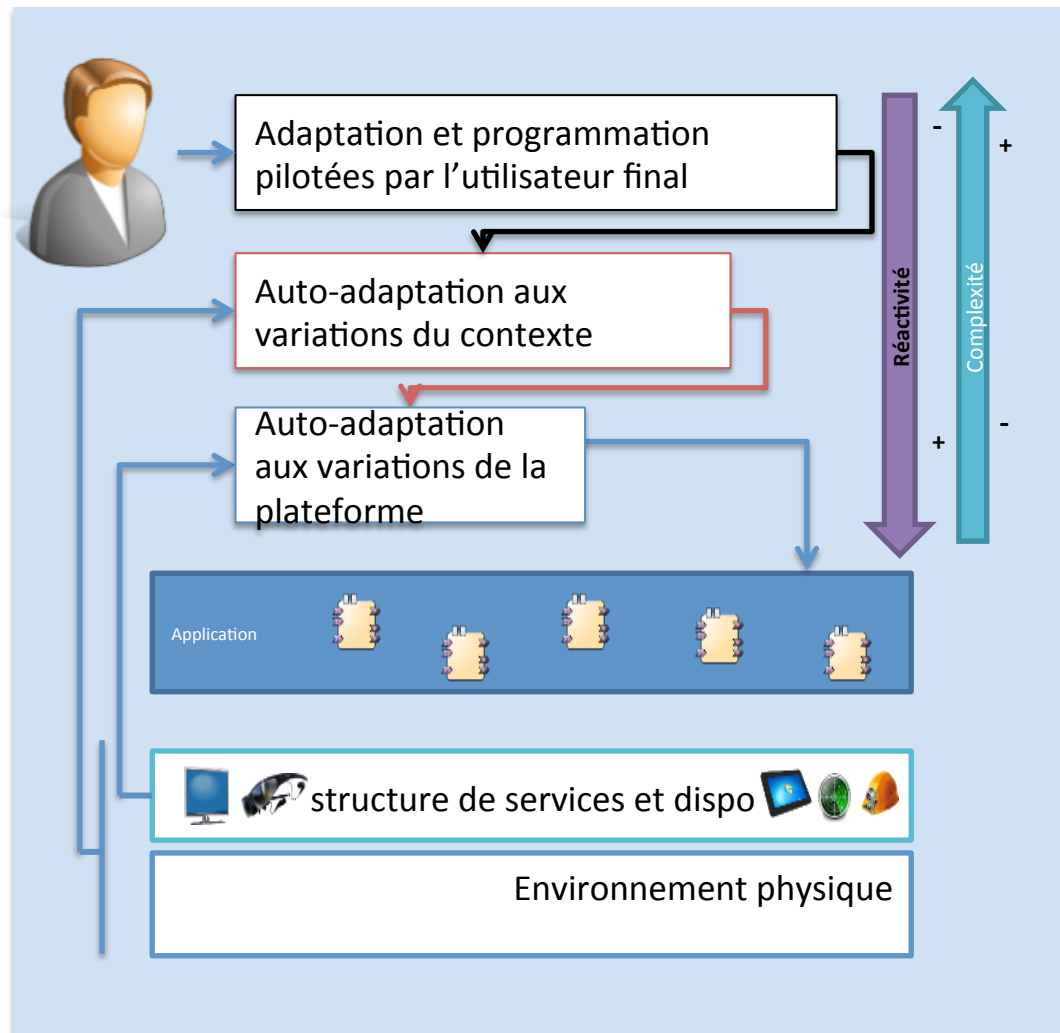
Adaptation dynamique de systèmes ambiants

[Tigli, Lavirotte, Riveill]

CONTINUUM

U-INSITHER

U-HEALTH



- Assurer la **Continuité de service** pour un utilisateur mobile dans un éco-système d'objets connectés

- Modèles** pour l'adaptation dynamique
- Expérimentations** avec utilisateurs (validation)

Moteur de recherche exploratoire [Marie, Legrand, Gandon, Palagi, Giboin]

Discovery Hub.



« J'aimerais découvrir des choses sur Claude Monet »

The screenshot shows a search result for 'Art rock' on the Discovery Hub platform. The page includes a search bar, navigation links like 'Add to Favorites' and 'Add to a Collection', and a detailed description of the term 'Art rock'. It also features a 'Stylistic Origin' section with links to 'Experimental rock' and 'Art music', and a 'YouTube's Videos For Art Rock And David Bowie' section with video thumbnails.

- Aider l'utilisateur à la recherche exploratoire (RE) en facilitant la navigation, la compréhension et le tri des résultats

The diagram shows a flow from a portrait of Claude Monet to a network graph with gears, representing the search process. The DBpedia logo is at the top. To the right, a list of four numbered steps is shown: 1, 2, 3, and 4, each with a corresponding input field. Below the list, the text reads: 'Results identification', 'Ranking', 'Sorting/categorization', and 'Explanation'. At the bottom, two mathematical formulas are provided:

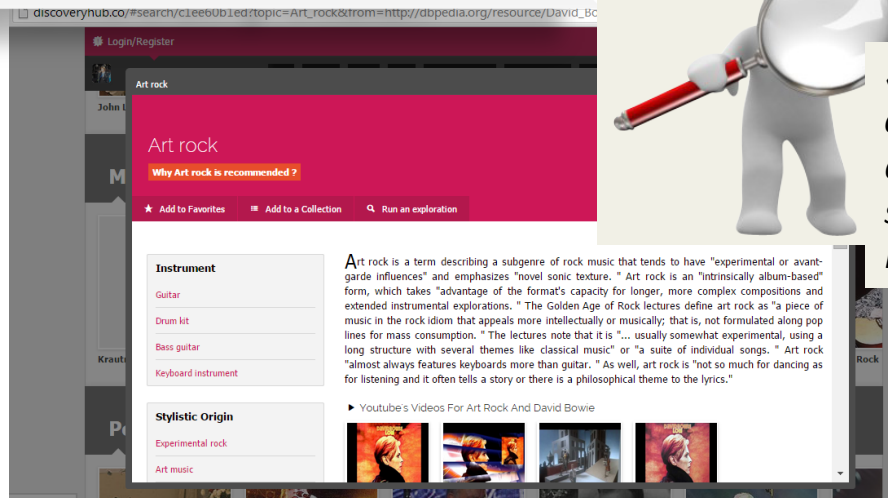
$$a(i, n + 1, o) = s(i, n, o) + \sum_j w(i, o) * \frac{a(j, n, o)}{\text{degree}_j}$$

$$a(i, n) = \prod_{o \in O} [a(i, n, o)] / \log(\text{degree}_i)$$

- Modélisation de l'utilisateur et de la tâche de RE
- Tri sémantique des résultats selon les préférences utilisateurs
- Explication des résultats (textes et graphes)
- Expérimentations avec utilisateurs (validation)

Moteur de recherche exploratoire [Marie, Legrand, Gandon, Palagi, Giboin,]

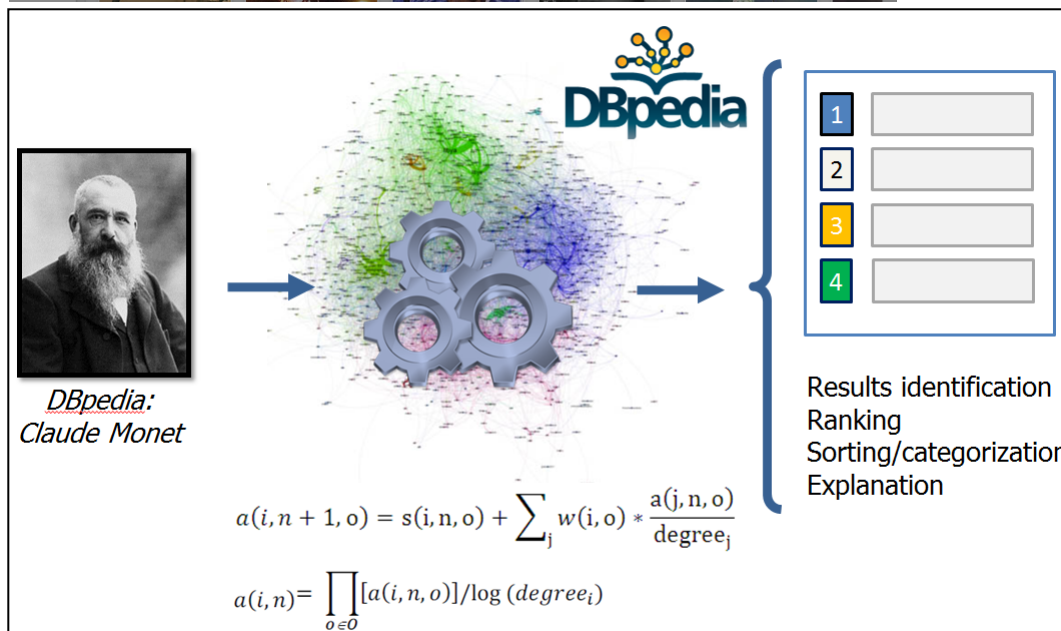
Discovery Hub.



« J'aimerais découvrir des choses sur Claude Monet »



- Aider l'utilisateur à la recherche exploratoire (RE) en facilitant la navigation, la compréhension et le tri des résultats



- Modélisation de l'utilisateur et de la tâche de RE
- Tri sémantique des résultats selon les préférences utilisateurs
- Explication des résultats (textes et graphes)
- Expérimentations avec utilisateurs (validation)

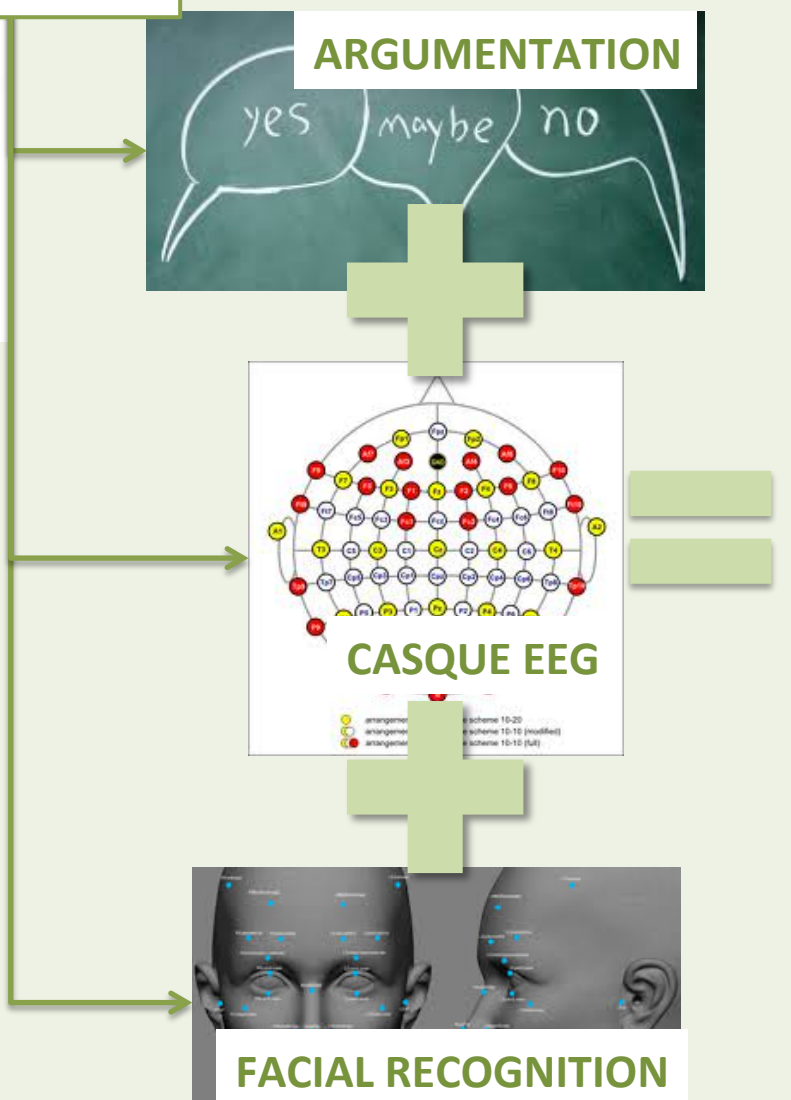
Emotions et Argumentation

[Villata, Cabrio]

SEEMPAD

Université  de Montréal

EXPERIENCE AVEC
LES UTILISATEURS



- Observer les **réactions émotionnelles** des personnes **argumentant** lors d'un débat sur un sujet donné

- Corrélations entre les **émotions** des participants et les **structures argumentatives** dans la discussion

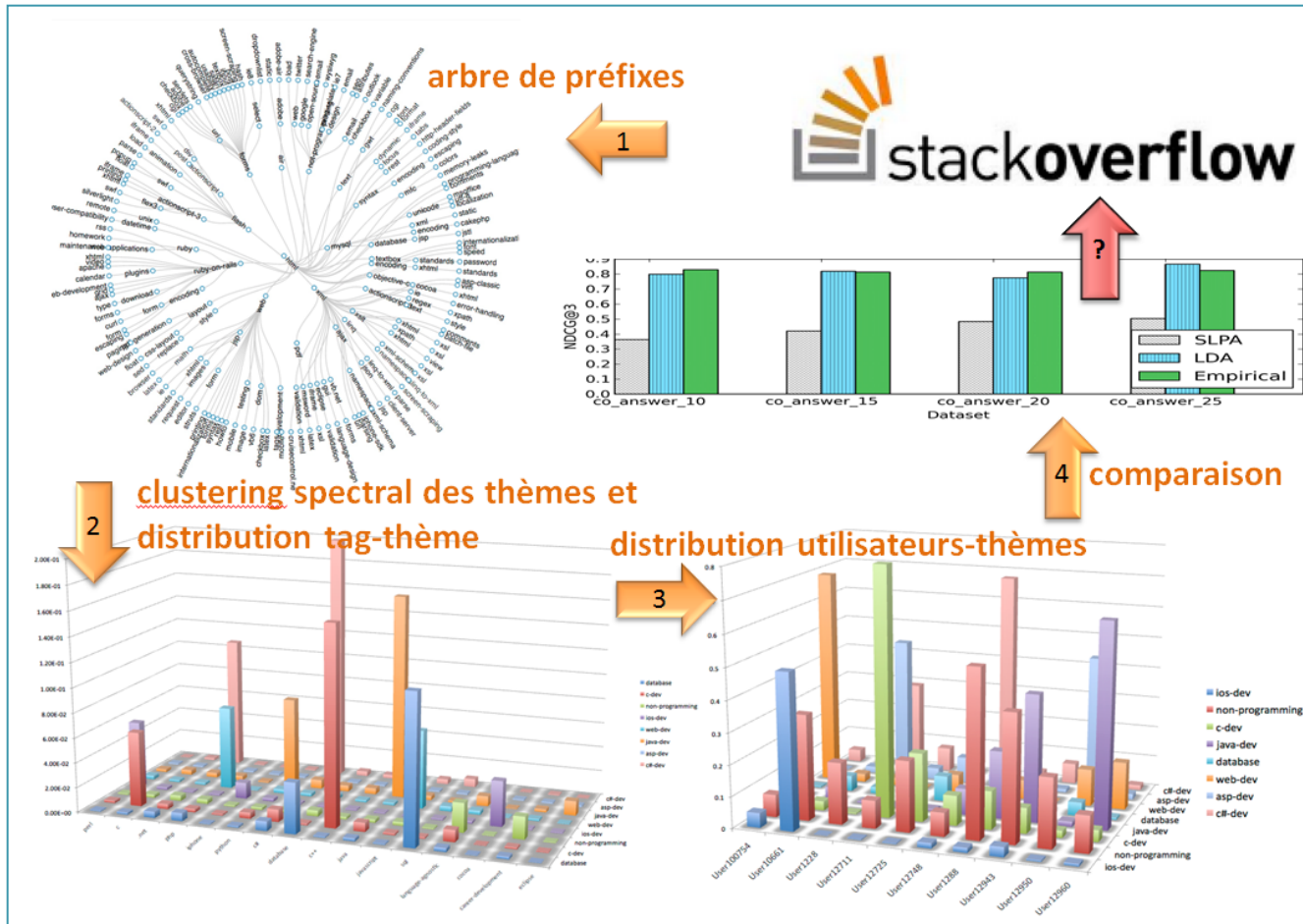
FORUM

FOrmalizing and Reasoning with Users and Models

interactions entre les utilisateurs

[Meng, Faron-Zucker, Gandon]

Analyse de communautés en ligne

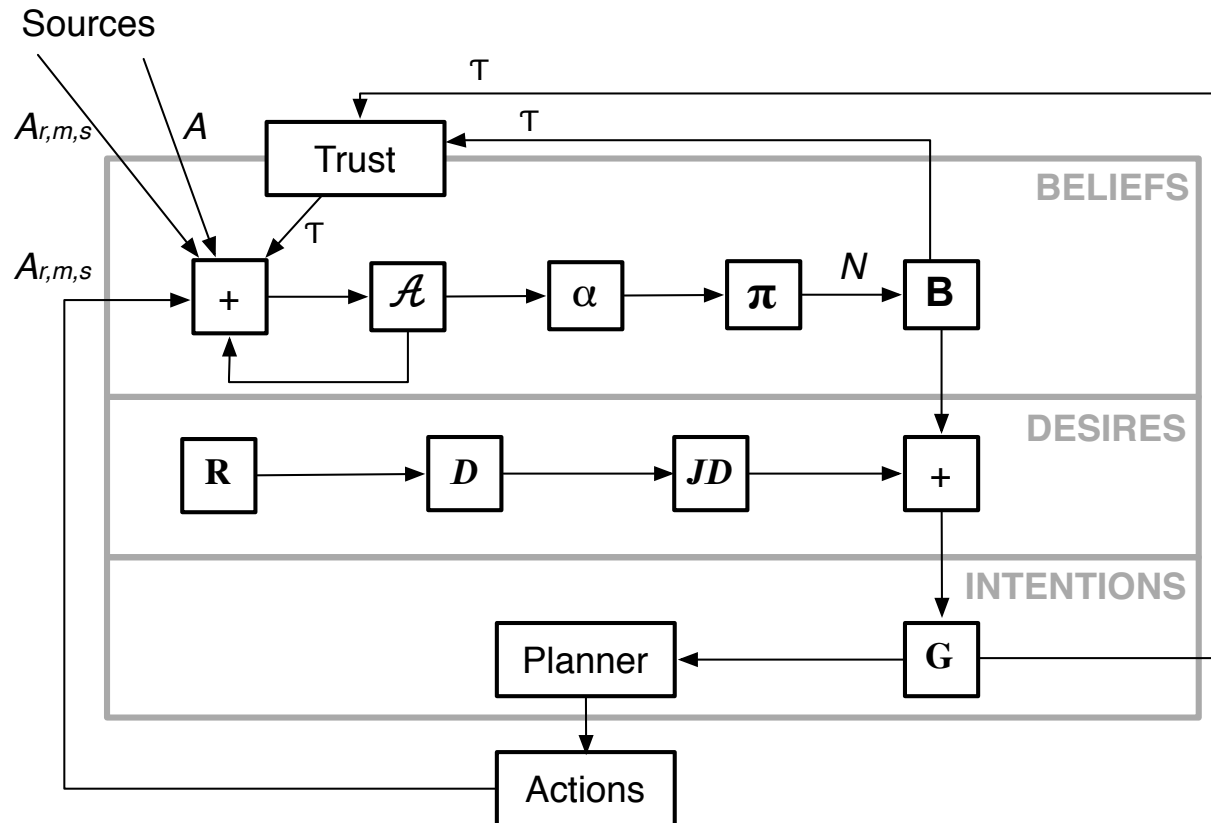


- Optimiser la productivité des recherches utilisateurs dans les forums en ligne

- Exploiter les spécificités de la structure du contenu de ces forums
- et de la communauté qui l'a produit

Modèles d'agents intelligents (BDI)

[Pereira, Tettamanzi, Villata]



- **Modéliser** le comportement intentionnel d'un agent

- Modélisation des croyances basée sur la **théorie des possibilités** et la **théorie de l'argumentation**
- Prise en compte de la **confiance** et de l'**incertitude** dans les sources d'information

FORUM

FOrmalizing and Reasoning with Users and Models

raisonnement avec des graphes

Moteur sémantique corese/kgram [Corby, Faron-Zucker]

Corese/KGRAM 3.1 - Wimmics INRIA I3S - 2015-03-30

File Edit Engine Debug Query Template Explain ?

System Query1 x +

Query Validate to SPIN to SPARQL Prove Trace Search Refresh stylesheet Default stylesheet

```

18 (count(?doc) as ?c)
19 (uri(?fa) as ?aa)
20 (uri(?fb) as ?bb)
21 (uri(concat(b:coauthor, "?", ?c) as ?r)
22 where {
23   ?doc dc:creator [rdfs:member ?a; rdfs:member ?b]
24   ?a rdfs:label ?fa
25   ?b rdfs:label ?fb
26   filter(?fa < ?fb)
27
28   optional { [rdfs:label ?fa ; b:member ?ta] . ?ta rdfs:label ?la }
29   optional { [rdfs:label ?fb ; b:member ?tb] . ?tb rdfs:label ?lb }
30
31 }
32 group by ?fa ?fb
having(?c >= 10)

```

Graph XML Validate

```

31 text-alignment:center;
32 size:17;
33 size-mode:fit;
34 fill-color:yellow;
35 shape:circle;
36 }
37
38 node.Class{
39   text-size:9;
40   text-color:black;
41   text-style:bold;
42   text-alignment:center;
43   size:17;
44   size-mode:fit;
45   fill-color:blue;
46   shape:circle;
47 }
48
49 edge{
50   text-color:black;
51   text-size:12;
52   size:2;
53   fill-color:grey;
54   text-alignment:center;
55   shape:cubic-curve;
56 }
57 node.keia {
58   text-size:12;

```

- Représentation des connaissances basée sur des graphes
- Recherche et Inférences dans les graphes
- Implémentation des standards du Web sémantique

- Semantic Web Factory: RDF/S, SPARQL 1.1 Query & Update, Inference Rules, Transformations
- Open Source CeCILL-C
- Knowledge Graph Abstract Machine avec 3 Proxies (Producer, Matcher, Evaluator)

ISICIL

Discovery Hub

Smilk

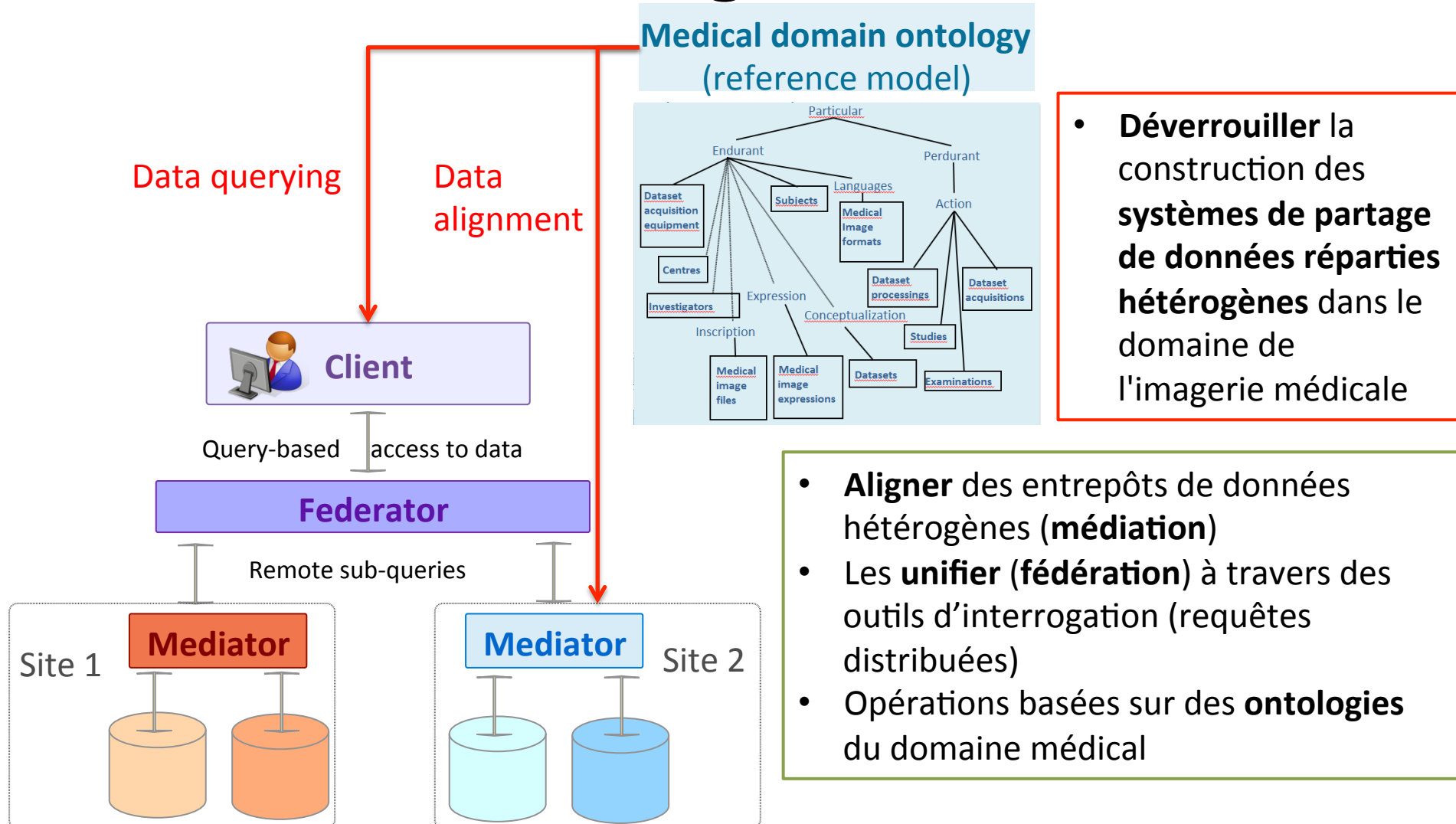
OntoCompo

SynchroBot

...

Fédération de données basée sur des ontologies

[Montagnat et al.]



FORUM : Perspectives

- **Modélisation** de toutes les dimensions de l'utilisateur (individu et communautés)
 - profil, préférences, contexte, relations sociales, croyances, émotions...
- **Prise en compte pour une adaptation du système à tous les niveaux**
 - recommandations personnalisées, explications, interfaces adaptées, adaptations dynamiques, analyse temporelle...
- **Modèles à base d'ontologies et de graphes**
 - **formalisation** de la sémantique sociale, intégration de connaissances hétérogènes