

# Automatic mappings discovery between ontologies in P2P context

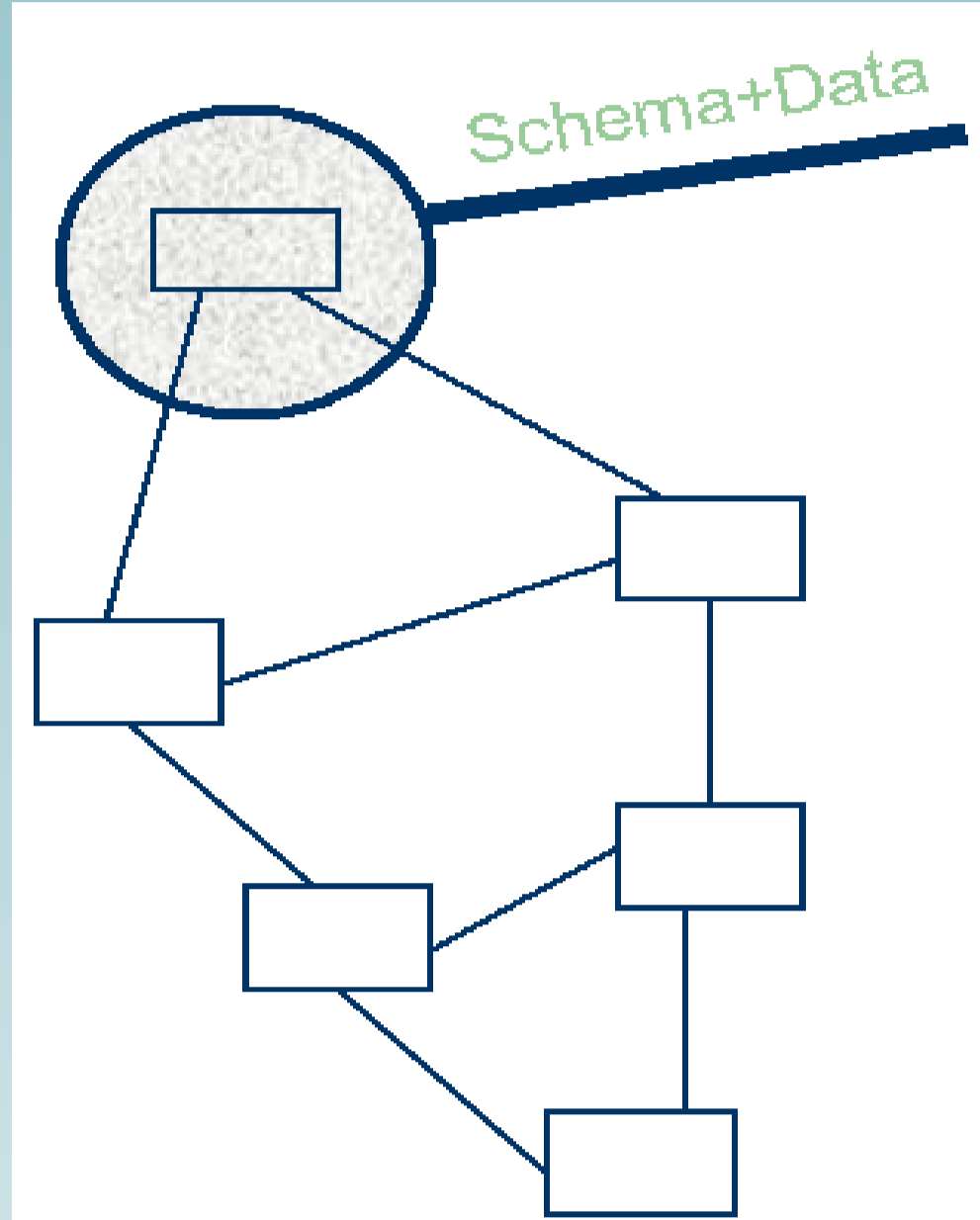
Rémi Tournaire – Supervisors : Marie-Christine Rousset & co : Jean-Marc Petit (LIRIS/BD)  
LIG/HADAS team – thesis started at september 2007

**Context: P2P network for sharing ressources semantically annotated by classes**

**Goal: automatically discover mappings between classes of different peers**

→ **By combining various methods: data mining, automatic reasoning**

## Issue



### Querying P2P network

Semantic-guided topology based on logical mappings between peers

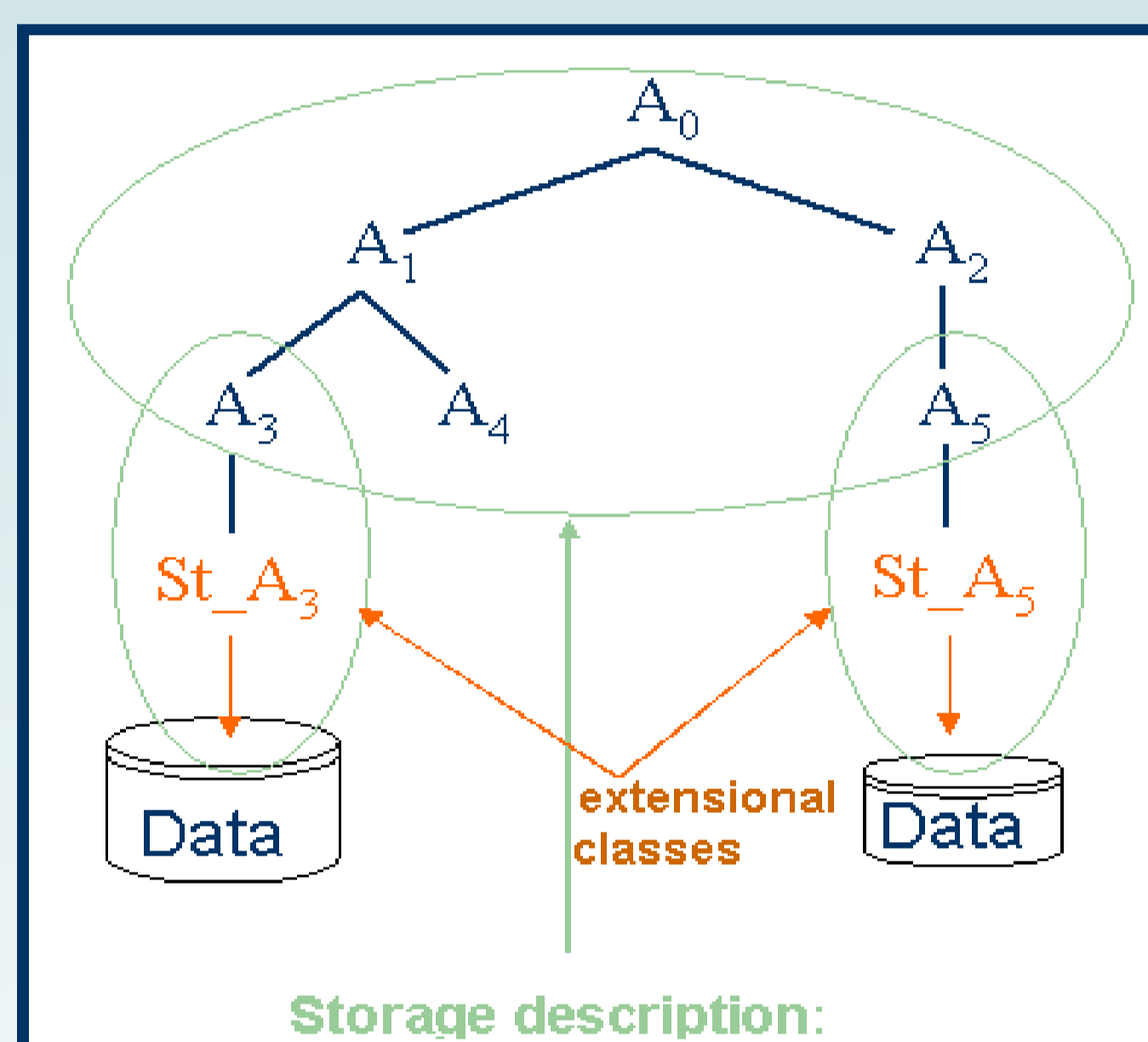
Somewhere: an existing scalable P2P platform

### Peer structure :

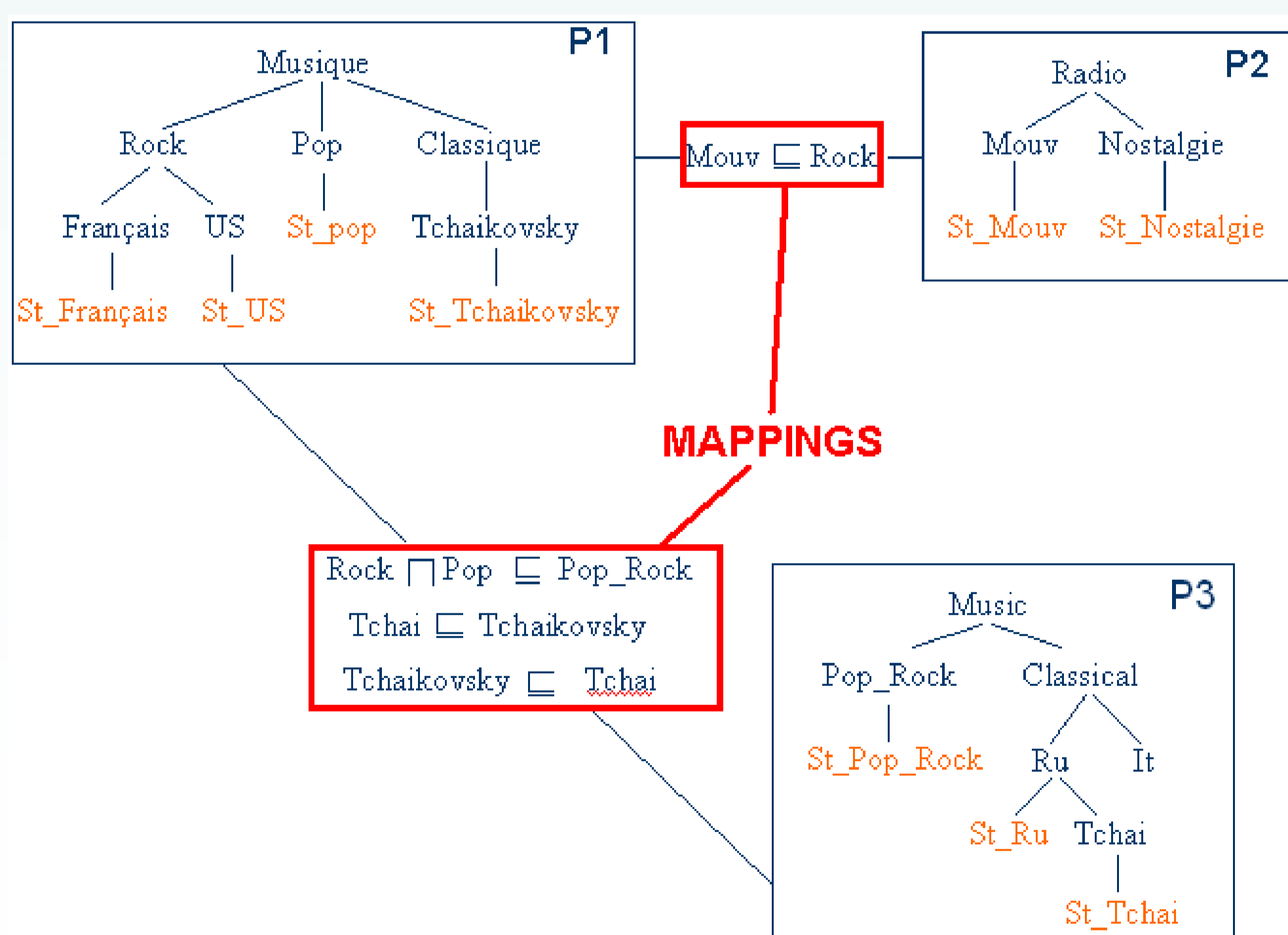
Hierarchical organization between **classes** and subclasses

**Ressources** referenced by unique identifier, annotated by classes

**Metadata** available and embedded in resources



### Mappings between peers



Links between ontologies

Now : manually declared

→ to be discovered automatically

## Realized and current work

### Modeling and Estimating mapping probabilities

- useful for selecting among candidate mappings
- different kind of modelisations to be compared

The probability of mapping  $m$  is conditioned by observations on resources related to  $m$

- bayesian estimation by data mining on metadata associated to these resources

Our model bridges logic and probability,

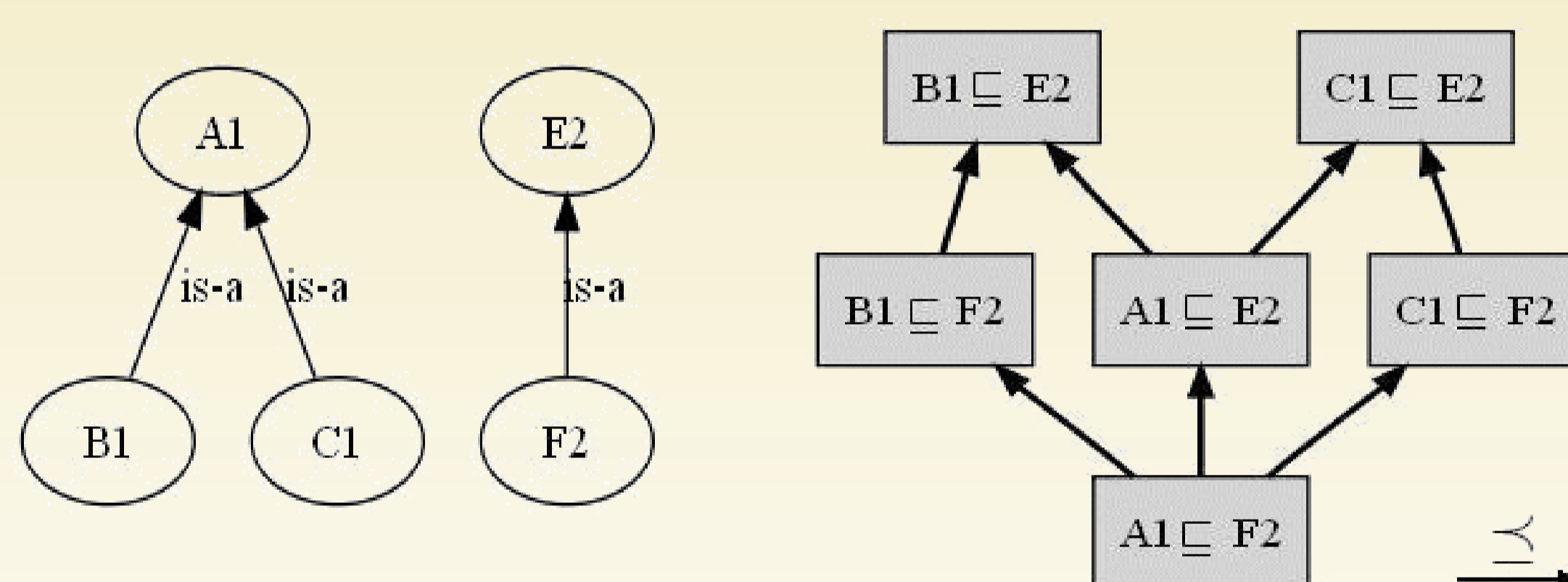
- under some *coherence assumption* between semantic class annotations and metadata

### Candidate mappings generation

Huge number of all the possible mappings in real cases

→ *Optimized enumeration algorithm*

- based on a pre-order on mappings exploiting semantics.



2 ontologies  $O1$  &  $O2$

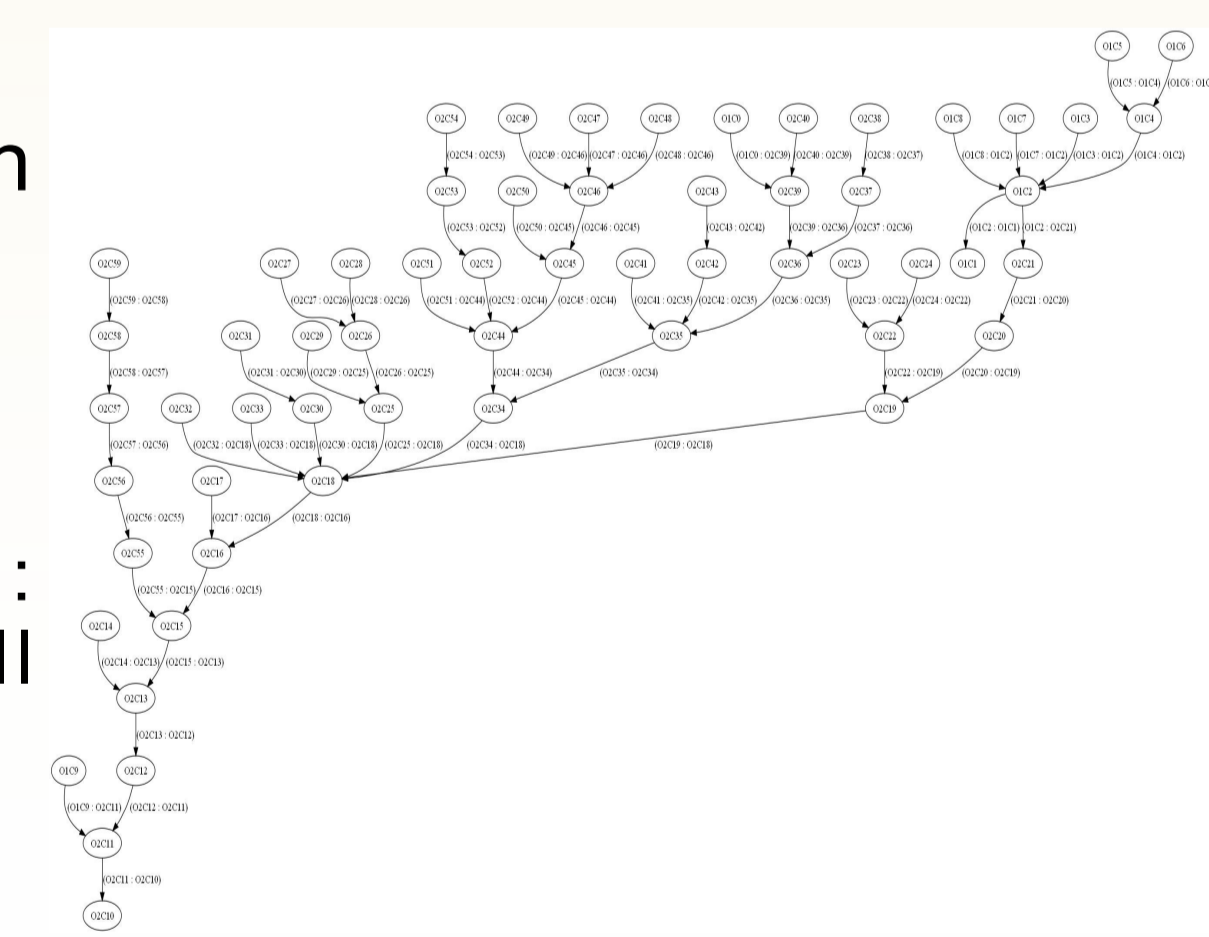
pre-order for candidate mappings from  $O1$  to  $O2$

Order built by automatic reasoning

### Experiments on artificial generated data (ontology & metadata):

- different versions of enumeration algorithm
- mapping probability estimation

Quantative & qualitative measurements : time complexity, precision & recall measures, noise robustness.



### Experiments on real data :

- directory set of Ontology Alignment Evaluation Initiative
- Yahoo and Google directories
- Collected taxonomies of different peers

### Setting and Funding:

WebIntelligence, Cluster ISLE Rhône-Alpes

Teaching at ESISAR, Grenoble INP

CLUSTERS DE RECHERCHE RHÔNE-ALPES



### Publication:

**IAF08-Journées Nationales de l'IA fondamentale**

Fouille de méta-données pour la découverte automatique de mappings entre taxonomies : une approche combinant logique et probabilités (R. Tournaire, M-C. Rousset)

**BDA09-Base de données avancées**

Découverte de mappings probabilistes entre taxonomies (R.Tournaire, J-M. Petit, M-C. Rousset, A.Termier)

