



# DBWeb: Research Topics

Pierre Senellart



November 22, 2010



# The DBWeb team

- The team:
  - 4 permanent researchers (+1 multi-affiliated member)
  - 6 PhD candidates (+2 multi-affiliated students)
  
- At the confluence of four areas of research:
  - Web data management, social networks
  - XML database systems
  - Database theory
  - Cognitive science



## Web data management, social networks

- Unsupervised **information extraction** from the structured Web (PhD thesis of Nora Derouiche)
- **Web archiving**: RSS feeds, deep Web, and other new forms of Web information (PhD thesis of Marilena Oita)
- **Trust** inference, propagation, and management in **social networks** (PhD theses of Silviu Maniu and Imen Ben Dhia)
- Large-scale **recommendation systems** (PhD thesis of Modou Gueye)
- **Data cleaning** and data enrichment using Web sources



# XML database systems

- A **probabilistic XML** system (PhD thesis of Asma Souihli)
- **XML diffs** and application to document integration
- **Access control** in XML documents
- **Pruning** of **XQuery** queries (see Talel's talk)



- Probabilistic XML (PhD thesis of Evgeny Kharlamov)
- Answering queries using views, in relational and XML settings
- Answering queries in the presence of limited access patterns, e.g., in the deep Web



- **Temporal information** in natural language: modeling, extraction (PhD thesis of Damien Munch)
- **Evolutionary** aspects of communication and language
- Relevance, surprise, coincidences: a **Kolmogorov complexity** explanation