



DES ENJEUX DE LA GESTION DE L'INFORMATION SUR LE WEB

le 8 avril 2014 15h30

ENS Rennes, Salle du conseil
[Plan d'accès](#)

Intervention de Pierre Bourhis (CNRS - LIFL - INRIA Lille)
Séminaire du département Informatique et télécommunications.



La recherche d'information a été depuis la création de l'informatique un domaine primordial. Elle a été pendant très longtemps séparée en trois domaines très différents :

La réponse précise à une question posée sur un ensemble de faits (Base de données)

La dérivation de nouvelles connaissances à partir d'un ensemble initial de connaissances et de règles de dérivations (Intelligence Artificielle)

L'apprentissage de relations entre informations non structurées (Data mining, Extraction d'information)

Si dans les années 70-90, le domaine ayant le plus succès a été les bases de données avec des entreprises multimilliardaires (Oracle, IBM...), avec l'explosion des données non structurées apparues sur le Web, les problèmes de Data Mining ont commencé à prendre une place très importante aujourd'hui. Dans cet exposé, nous présenterons deux points importants: tout d'abord, les frontières entre les différents domaines ne sont plus hermétiques; ensuite le domaine des bases de données est toujours aussi vivant et les challenges apportés par le Web sont nouveaux et très importants. Après avoir présenté brièvement les notions initiales des bases de données: homomorphisme, complexité combinée, complexité dans les données, notion d'optimisation de requêtes; nous présenterons les différents enjeux pour gérer de l'information sur le Web: hétérogénéité des données, distribution des données, explosion de la quantité des données. Pour illustrer ces problèmes, nous présenterons RDF qui est le standard W3C pour la représentation et la diffusion de données sur Web. Ce langage est l'exemple classique du mariage des problèmes classiques de bases de données et de dérivation de nouvelles connaissances. Nous présenterons quelques résultats dans ce domaine et nous conclurons par comment intégrer dans ce contexte l'incertitude dans les données résultant de l'apprentissage de relations entre elles.

THÉMATIQUE(S)

Formation, Recherche - Valorisation

CONTACT

François Schwarzenruber

Mise à jour le 9 septembre 2019

CONTACT

Raphaël Truffet

ARCHIVES

[Séminaires 2020-2021](#)
[Séminaires 2019-2020](#)
[Séminaires 2018-2019](#)
[Séminaires 2017-2018](#)
[Séminaires 2016-2017](#)
[Séminaires 2015-2016](#)
[Séminaires 2014-2015](#)
[Séminaires 2013-2014](#)
[Séminaires 2012-2013](#)
[Séminaires 2011-2012](#)
[Séminaires 2010-2011](#)
[Séminaires 2009-2010](#)
[Séminaires 2008-2009](#)
[Séminaires 2007-2008](#)
[Séminaires 2006-2007](#)
[Séminaires 2005-2006](#)
[Séminaires 2004-2005](#)
[Séminaires 2003-2004](#)
[Séminaires 2002-2003](#)