

**eric** ENTREPÔTS, REPRÉSENTATION  
 et INGÉNIERIE des CONNAISSANCES

**dis** Département d'Informatique et de Statistique  
 Institut de la Communication

**Master 2 ECD**  
 2015-2016

**Entreposage et fouille de données complexes**

Sabine Loudcher  
 Jérôme Darmont

UNIVERSITÉ LYON 1  
 UNIVERSITÉ LYON 2  
 UNIVERSITÉ DE LYON

**Contexte**

**Nature des données actuelles**

- Les données bien maîtrisées sont « simples »
  - Données numériques
  - Données symboliques
- Besoin de traiter et d'analyser des données plus diverses et hétérogènes
  - Données multimédia
  - Données du Web

**eric** Entreposage et fouille de données complexes – M2 ECD 2015-2016

**Contexte**

**Une définition des données complexes**

Multiformats (orientée syntaxe)  
 OU  
 Multistuctures  
 OU  
 Multisources  
 Au moins deux caractéristiques  
 OU  
 Multimodales  
 OU  
 Multiversions

**eric** Entreposage et fouille de données complexes – M2 ECD 2015-2016

**Contexte**

**Données multiformats**

**eric** Entreposage et fouille de données complexes – M2 ECD 2015-2016

**Contexte**

**Données multistuctures**

**SQL**

**Not Only SQL**

**RDF**

**eric** Entreposage et fouille de données complexes – M2 ECD 2015-2016

**Contexte**

**Données multisources**

**eric** Entreposage et fouille de données complexes – M2 ECD 2015-2016

### Contexte

#### Données multimodales

Entreposage et fouille de données complexes – M2 ECD 2015-2016

### Contexte

#### Données multiversions

Traditional setting	Multi-version setting
<pre> &lt;last_modified on="2010-01-01"/&gt; &lt;element_x&gt;   &lt;applies to="T1.X"/&gt;   &lt;OLD CONTENTS* &lt;/element_x&gt;           </pre> <p>Modification on 2011-11-01</p>	<pre> &lt;element_x&gt;   &lt;version number="1"/&gt;   &lt;peristence&gt;     &lt;valid from="2010-01-01" to="2011-10-31"/&gt;     &lt;applies to="T1.X"/&gt;   &lt;/peristence&gt;   &lt;OLD CONTENTS* &lt;/version&gt;   &lt;version number="2"/&gt;   &lt;peristence&gt;     &lt;valid from="2011-11-01" to="9999-99-99"/&gt;     &lt;applies to="T1.X; T2.YZ"/&gt;   &lt;/peristence&gt;   &lt;NEW CONTENTS* &lt;/version_x&gt;           </pre>

Web Page Archives

Entreposage et fouille de données complexes – M2 ECD 2015-2016

### Contexte

#### Et pour plus d'axes de complexité...

Entreposage et fouille de données complexes – M2 ECD 2015-2016

### Contexte

#### Des données complexes aux mégadonnées

www.libbigdatahub.com

Entreposage et fouille de données complexes – M2 ECD 2015-2016

### Contexte

#### Des données complexes à la décision

eric.ish-lyon.cnrs.fr

Entreposage et fouille de données complexes – M2 ECD 2015-2016

### Thèmes abordés

#### OLAP sémantique (S. Loudcher)

- Données textuelles
  - ETL pour le texte
  - Cubes de textes
  - Agrégation textuelle
- Opérateurs OLAP
  - structuration
  - prédiction
  - explication
  - agrégation sémantique
- Théorie des graphes
  - Réseaux d'information
  - Cubes de graphes
  - Graphes avec des cubes
  - Opérateurs topologiques

Entreposage et fouille de données complexes – M2 ECD 2015-2016

### Thèmes abordés

■ Décisionnel dans le nuage (J. Darmont)

The diagram illustrates a user on the left connected to a cloud environment. The cloud contains a 'Fournisseur d'accès' (Access Provider) and 'ED' (Data) represented by a cylinder and a padlock. Arrows labeled 'Transfert réseau' (Network Transfer) show data moving from the cloud to the user and from the user to the cloud. A small car icon is at the bottom of the cloud.

e<sup>2</sup>iC Entreposage et fouille de données complexes – M2 ECD 2015-2016 13

### Méthodologie

■ Travail sur articles scientifiques

	OLAP sémantique	Décisionnel dans le nuage
Présentation du thème	12/01	5/01
Etude d'article et de présentation de recherche	19/01	7/01
Synthèse d'articles	26/01	15/01
Évaluation	12/02	

e<sup>2</sup>iC Entreposage et fouille de données complexes – M2 ECD 2015-2016 14

### Méthodologie

■ Évaluation

- Organisation
  - Présentation par 3 ou 4 étudiants selon le site
  - Étude de 2 ou 3 articles (possibilité de rajouter des articles)
  - Synthèse bibliographique
  - Présentation orale le 11/02/2016
  - Contrôle continu, pas de rattrapage possible
- Contenu de la présentation
  - Thématique commune des articles : contexte et problématique du thème
  - Résumer les articles : objectif, méthodes ou solutions et résultats
  - Définir des critères de comparaison
  - Faire une comparaison
  - Discussion/conclusion

e<sup>2</sup>iC Entreposage et fouille de données complexes – M2 ECD 2015-2016 15

sabine.loudcher@univ-lyon2.fr

jerome.darmont@univ-lyon2.fr

e<sup>2</sup>iC Entreposage et fouille de données complexes – M2 ECD 2015-2016 16