



**HAL**  
open science

## Portail d'université : proposition d'un service OLAP via l'espace numérique de travail.

Frédérique Peguiron, Odile Thiéry

### ► To cite this version:

Frédérique Peguiron, Odile Thiéry. Portail d'université : proposition d'un service OLAP via l'espace numérique de travail.. SEMINAIRE INTERNATIONAL SUR LES PORTAILS D'ENTREPRISE, Algérie Telecom, Jun 2007, Alger. inria-00155552

**HAL Id: inria-00155552**

**<https://inria.hal.science/inria-00155552v1>**

Submitted on 18 Jun 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Portail d'université : proposition d'un service OLAP via l'espace numérique de travail.**

Frédérique Peguiron, Odile Thiery

LORIA, Campus Scientifique, B.P. 239,  
54506 Vandoeuvre-les-Nancy Cédex, France  
{Peguiron, Thiery,}@loria.fr  
<http://www.loria.fr/~peguiron>

## **Résumé.**

Le transfert d'un système d'information en système d'information décisionnel repose sur les bases métiers orientées vers les acteurs de l'université par la prise en compte de la modélisation des utilisateurs. Le développement d'un portail de l'université tient compte de l'évolutivité du contexte de l'université et de la prise en compte des systèmes d'information alternatifs. La description des ressources électroniques et la modélisation des acteurs montrent aujourd'hui la complémentarité de deux univers : les mondes de l'indexation et les mondes du décisionnel reliés par les entrepôts de données. A partir d'une expérimentation pour aider l'acteur «enseignant-chercheur» à enrichir un état de l'art par une fouille de données à partir d'un entrepôt de données, nous montrons les possibilités d'exploitation de données qui font partie de la partie immergée d'un système d'information universitaire. Nous portons les résultats de l'expérimentation au travers d'une application développée à partir d'un produit en open source Openi pour enrichir un espace numérique de travail d'un service d'analyses multidimensionnelles adaptées aux différents acteurs du portail de l'université.

## **Mots-clés**

Portail universitaire, Espace Numérique de Travail, bibliométrie multidimensionnelle, analyse multidimensionnelle, entrepôt de données, OLAP.

## 1. Introduction

Actuellement nous assistons à des développements rapides des Environnements Numériques de Travail dans plusieurs universités qui ont fait émerger des difficultés relatives à la conduite du changement et une implication des différents acteurs de l'université. Les Systèmes d'Information sont au cœur des organisations. Les universités se sont organisées autour de consortium pour rayonner en Université Numérique de Région ou Université Numérique Thématique pour proposer des Espaces Numériques de Travail (ENT) à leurs utilisateurs. Ces ENT ne sont pas seulement une juxtaposition d'outils, mais proposent des services pour permettre une entrée pédagogique de ses acteurs. Après avoir rappelé les principaux services mis à la disposition des utilisateurs qui représentent la partie émergée du système d'information, nous proposons de nous pencher sur les données de la partie immergée du système d'information pour lui permettre d'atteindre la dimension d'un système d'information stratégique au bénéfice d'une facilitation de gestion de l'organisation par ses acteurs.

A l'instar des pôles de compétitivité en Entreprise naissent les pôles de recherche et d'enseignement supérieur en Université qui font office de levier dans le développement des universités à l'échelon du territoire voire à un échelon mondial. Nous voyons bien là de nouvelles perspectives d'analyse pour mettre en œuvre un système d'observatoire au service d'une nouvelle gouvernance des universités.

## 2. Enjeux

Nous avons remarqué plusieurs enjeux qui vont en faveur de la spécification de notre problématique. Les étudiants ont pour but de réussir, apprendre mieux, autrement, efficacement par la construction d'un projet professionnel qui doit les aider à trouver un emploi. Les étudiants, au moment de prendre une inscription dans un établissement opèrent selon une démarche qualité. L'enseignant passe du rôle d'enseignant à celui d'auteur. Il s'agit de corrélérer les besoins des étudiants et les besoins des enseignants. Les administrations gagnent à diffuser et partager les informations pour résoudre des problèmes, échanger des points de vue, reproduire ou innover.

Si nous analysons les gouvernances des entreprises et des universités, c'est dans la recherche de la satisfaction du **client/utilisateur** que nous pouvons faire un parallèle entre ces deux mondes. Pour les entreprises la démarche de gouvernance est à but lucratif, pour les universités la démarche de gouvernance concerne davantage un positionnement et une visibilité de l'organisation, le but qui peut tendre également vers une recherche de rentabilité, vise principalement à une facilitation de la gestion. L'entreprise est en quête d'un positionnement de performance au niveau de son capital, l'objectif d'une université est d'atteindre le label niveau 1 de qualité dans le classement des universités. Dans le premier cas cela passe par la satisfaction des **clients**. Dans le second cas, le but est de satisfaire les **utilisateurs**. La satisfaction des clients en entreprise se formalise en termes de coûts. La satisfaction des utilisateurs en université s'illustre par la satisfaction de leurs besoins. Un système d'information stratégique repose sur des outils d'analyse, principalement l'entrepôt de données dans notre cas, qui permet d'évaluer les «prises de risques» ou au contraire les «bénéfices ramenés» à prendre en compte les enjeux pour l'urbanisation d'un système d'information universitaire.

### 3. Les technologies employées pour notre contribution

Rappelons qu'un ENT est un dispositif qui permet à tous les usagers de disposer des services en rapport avec leur activité à partir d'un point unique, sa conception implique une réflexion globale sur l'organisation en tenant compte des acteurs et des objectifs. Le Système d'Information de l'Université est complexe et hétérogène. Résumons de façon non exhaustive les services classiquement proposés dans un ENT qui constituent une entrée pédagogique pour les acteurs et représentent la partie émergée du Système d'Information :

- Courriel, forum, agenda, Plan de travail,
- Podcasting<sup>1</sup> : fichiers MP3, Mpeg,
- Portefeuille de compétences ou Portfolio : porte document partageable, flux RSS,
- Banques d'images et d'animations,
- Content Management System (CMS) ou système de gestion de contenu,
- Volet pédagogique : cours, exercices,
- Volet documentaire : bases de connaissances.

L'analyse des compétences des acteurs sur un système d'information montre que la recherche d'information est une suite de processus où toutes les étapes sont en interaction. Par la dissociation des fonctions, des besoins, des activités par type d'acteurs nous raisonnons selon des niveaux par rapport aux usages des informations. Trois niveaux sont concernés ici : le niveau acteur, le niveau administratif et le niveau enseignement.

- Le niveau acteur permet une première typologie des acteurs autour de 3 classes, qui fait apparaître des étudiants, des enseignants et des administratifs.
- Le niveau enseignement permet d'identifier des bases «référénts» corrélées avec les acteurs précédemment identifiés : des bases de cours plutôt destinées aux étudiants, des bases de références au service des enseignants et des bases de textes réglementaires à la destination des administratifs.
- Le niveau administratif recense des données relatives à la situation administrative de l'acteur étudiant, des données relatives à la situation administrative de l'acteur enseignant et des données de gestion administrative et financière des étudiants, des enseignants et des formations utiles à l'acteur administratif.

Le scénario qui suit tient compte des différents niveaux des acteurs pour tirer parti de bases de connaissances. Nous montrons à partir d'une phase expérimentale la faisabilité d'effectuer une analyse bibliométrique multidimensionnelle à l'aide d'un entrepôt de données au service d'un état de l'art pour l'acteur enseignant-chercheur. Après quoi nous avons tenu à développer un prototype en open source dans l'esprit de l'Internet 2e génération qui favorise la création de contenus numériques (Rosnay, 2006) par les usagers user generated contents.

#### 3.1 Analyse bibliométrique multidimensionnelle : phase expérimentale

Nous proposons par exemple à l'acteur «enseignant-chercheur» de compléter un état de l'art par l'analyse de listes de diffusion à partir d'un entrepôt de données. D'après Franco l'architecture de l'entrepôt de données comporte trois niveaux fonctionnels essentiels : le niveau acquisition des données, le niveau stockage des données et le niveau analyse de données (Franco, 1997). L'entrepôt de données doit intégrer les données les unes avec les autres afin d'assurer une cohérence sémantique globale. Il se compose d'un data warehouse, de

---

<sup>1</sup> Podcasting (un terme composé autour des mots iPod, webcasting, et broadcasting) est une technique qui permet de transférer et d'écouter automatiquement sur son baladeur MP3 les programmes audio d'un site, sans avoir à le visiter.

bases de données multidimensionnelles ou hypercubes et d'un ensemble d'outils permettant l'alimentation du data warehouse, son interrogation et la production de rapports, l'extraction intelligente des données par techniques de data mining enfin l'analyse décisionnelle. Nous restituons ici une méthode adoptée pour faire ressortir les tendances émergentes propre au thème de l'intelligence économique entre 2001 et 2005.

Nous proposons une méthode pour prendre connaissance d'un vocabulaire existant autour d'une thématique, de son évolution et de son contexte. Le thème retenu «l'intelligence économique» est un processus qui couvre plusieurs champs disciplinaires de façon transversale. Il s'agit de cerner les différents concepts propres à cette thématique. Cette façon de procéder permet de repérer les tendances émergentes, les acteurs, les réseaux, les parutions d'ouvrages et les conférences ou colloques en rapport avec ce thème. Concrètement nous sommes abonnées à des listes de diffusion et à des groupes de discussion autour : des outils de recherche d'information, des moteurs de recherche, de l'intelligence économique, de la gestion des connaissances, des outils de veille, des outils spécifiques à la documentation, des outils spécifiques aux bibliothèques.

| Liste de diffusion  | Thèmes  |
|---|---|
| <a href="mailto:adbs-info@cru.fr">adbs-info@cru.fr</a>  | L'association des professionnels de l'information et de la documentation en 1994, a pour objectif de faciliter les échanges d'informations, d'idées et d'expériences.   |
| <a href="mailto:adest@grenet.fr">adest@grenet.fr</a>  | Bibliométrie, scientométrie, infométrie, recherche théorique et appliquée.  |
| <a href="mailto:agents@yahoogroupes.fr">agents@yahoogroupes.fr</a>  | Consacrée aux agents intelligents. Imbriqué au site AgentLand.fr, cet espace permet d'échanger des solutions pour mieux maîtriser les agents, poser des questions, suggérer des améliorations, donner des avis sur un agent.  |
| <a href="mailto:biblio-fr@cru.fr">biblio-fr@cru.fr</a>  | Cette liste de diffusion regroupe bibliothécaires et documentalistes francophones, et toute personne intéressée par la diffusion électronique de l'information documentaire.  |
| <a href="mailto:cybercrise@yahoogroupes.fr">cybercrise@yahoogroupes.fr</a>  | Ce groupe de discussion est destiné à échanger et à faire évoluer la réflexion sur la gestion de crise.   |
| <a href="mailto:gredec@grenet.fr">gredec@grenet.fr</a>  | Liste dédiée à la mesure des sciences et techniques   |
| <a href="mailto:ienetwork@yahoogroupes.fr">ienetwork@yahoogroupes.fr</a>  | Ce groupe a pour objectif de regrouper les nouveaux et anciens étudiants, les professionnels de l'intelligence économique, promouvoir notre profession, échanger des offres d'emploi et des informations sur l'actualité de l'intelligence économique au niveau international.                                    |
| <a href="mailto:i-KM@yahoogroupes.fr">i-KM@yahoogroupes.fr</a> puis<br><a href="mailto:i-KMFORUM@yahoogroupes.fr">i-KMFORUM@yahoogroupes.fr</a> | Forums et listes de discussion du secteur de l'information-documentation  |
| <a href="mailto:intelligence-economique@yahoogroupes.fr">intelligence-economique@yahoogroupes.fr</a>  | Ce groupe de discussion est consacré à l'intelligence économique au sens large, c'est à dire la gestion de l'information externe : mise en place d'un système de veille, outils et méthodes, les aspects de protection de l'information, de renseignement, de benchmarking, d'influence, de knowledge management. |
| <a href="mailto:miste-esiee@yahoogroupes.fr">miste-esiee@yahoogroupes.fr</a>  | Ce forum dédié à la créativité et à l'utilisation des cartes heuristiques, cartes d'organisation d'idées, topogrammes, arbres à sens, schémas arborescents, cartes mentales et autres «mind maps»   |
| <a href="mailto:motrech@yahoogroupes.fr">motrech@yahoogroupes.fr</a>  | Cette liste, consacrée aux moteurs de recherche sur Internet, est un lieu d'échanges sur les problématiques, techniques, développements et évaluations/comparaisons des outils de recherche d'information sur Internet.   |
| <a href="mailto:netkm@egroups.fr">netkm@egroups.fr</a> puis <a href="mailto:netkm@yahoogroupes.fr">netkm@yahoogroupes.fr</a>                    | Club du Knowledge Management et de l'Intelligence Economique.   |
| <a href="mailto:newsletter@afnet.fr">newsletter@afnet.fr</a>  | Liste de diffusion de l'AFNeT   |
| <a href="mailto:prospective@egroups.fr">prospective@egroups.fr</a>  | Prospective sur Internet. De quelle manière Internet peut être un excellent outil pour détecter les nouvelles tendances, constituer un réseau d'experts, identifier les réseaux de collaborations.  |
| <a href="mailto:veille@egroups.com">veille@egroups.com</a> puis <a href="mailto:veille@yahoogroupes.fr">veille@yahoogroupes.fr</a>              | Liste consacrée aux thématiques de veille sur Internet. Historiquement, il s'agit de la première mailing-list française sur l'intelligence économique et stratégique sur Internet (1998).   |

Figure 1 : Listes de diffusion et groupes de discussion étudiés de 2001 à 2005

### a) Récapitulatif des groupes et listes comme support d'analyses

Ces listes de diffusion et groupes de discussion nous offrent un corpus de 724 messages après une équation de recherche autour du mot «colloque» dont nous ne retenons que les messages ayant trait aux événements concernant notre sujet de recherche (l'intelligence économique) via les événements qui s'y rapportent. Un premier tableau analytique est réalisé après dépouillement des listes de diffusion et groupes de discussion consacrés à l'intelligence économique – il permet la réalisation de tableaux synthétiques autour d'indicateurs en vue d'une fouille de données. Le tableau analytique est construit autour des rubriques suivantes : «**Événement**», «**Date**», «**Lieu**», «**Site dédié**», «**Organisateur**» et «**Objectifs**». Une simple lecture de ce tableau dense en quantité d'informations ne permet pas de faire une analyse fine et de mettre l'accent sur une évolution des événements. C'est pourquoi une seconde lecture dite «intelligente» permet de proposer un second tableau simplifié où sont rajoutées des rubriques pour chaque événement : Thèmes, Secteur, Type d'organisateur et Spécialité de l'organisateur. Par ces nouvelles rubriques nous mettons en place des indicateurs qui constituent des clés de lecture en vue d'une fouille de données pour mettre l'accent sur des émergences, des évolutions ou encore des tendances. La rubrique «**Thèmes**» concerne au principal thème abordé lors de l'événement. Le champ «**Secteur**» identifie le secteur global touché ciblé par l'événement. «**Organisateur\_type**» type l'organisateur par rapport à sa raison sociale relative à notre sujet de recherche.

### b) Préalable à l'analyse multidimensionnelle

Nous obtenons à l'aide d'Excel plusieurs tableaux pour chaque année de 2001 à 2005 dont nous restituons pour exemple le tableau relatif à l'année 2004 :

| Evenements             | Thèmes          | Secteur    | Objectifs  | Date | Organisateur type | Organisateur specialite     |
|------------------------|-----------------|------------|--|------|-------------------|-----------------------------|
| EGC                    | méthodes        | km         | extraction de connaissances à partir de données                    | 2004 | institut          | information                 |
| journées veille        | veille          | entreprise | recherche d'information  | 2004 | université        | Information communication   |
| journées linguistiques | outils-méthodes | km         | gestion des contenus   | 2004 | association       | Information économie        |
| VSST                   | méthodes        | km         | exploitation efficace des grandes masses de documents              | 2004 | université        | Informatique bibliométrique |
| congré                 | projet          | entreprise | compétitivité et innovation  | 2004 | association       | information                 |
| journées ADBS          | Exploitation    | Doc        | weblogs dans la publication et diffusion de l'information : enjeux | 2004 | association       | Documentation information   |

**Figure 2 : Tableau synthétique autour d'indicateurs pour l'année 2004**

### c) Analyse multidimensionnelle en vue d'une fouille de données

A partir de ce type de tableau, les fonctionnalités d'Excel, hormis les fonctions de calcul, de tri et de représentation graphique, ne nous permettent pas de mettre en valeur des tendances. Ayant à notre disposition le logiciel Cognos utilisé plutôt pour des analyses financières, nous l'exploitons pour procéder à une fouille de données pour faire émerger visuellement des dates clés en rapport avec les événements autour de l'intelligence économique et mesurer l'apparition de concepts. Cognos est un logiciel qui permet l'exploration d'une base

multidimensionnelle. Cognos regroupe des outils dont Powerplay<sup>2</sup> et Transformer<sup>3</sup>. Les tableaux de 2001 à 2005 réalisés à l'aide d'Excel sont importés dans une table Access. Nous utilisons le logiciel Transformer qui permet de créer un modèle où apparaissent les sources de données, les mesures, les cubes ainsi que la grille des dimensions. Nous pouvons afficher un diagramme qu'il est facile d'interpréter visuellement. On s'aperçoit ainsi que les universités, au cours des années, représentent régulièrement le nombre le plus important d'organisateur. Le secteur de la documentation est largement concerné depuis 2003.

A partir des résultats de cette première phase expérimentale d'analyse bibliométrique multidimensionnelle, nous avons pu mettre en perspective les thèmes abordés de l'état de l'art regroupés autour de quatre concepts qui sont : les tendances (idées, expériences, gestion de crise), l'information (actualité, national, international, protection), les outils (nouveaux outils, réseaux d'experts, méthodes) et la représentation (traitement des données, cartes heuristiques, cartes d'organisation d'idées) autour de l'intelligence économique – intelligence économique à forte connotation «gestion des connaissances». Deux verbes d'action sous tendent les relations entre concepts et thèmes : collaborer d'abord, puis partager.

#### d) Enrichissement des services de l'ENT

Au vu des bénéfices ramenés à utiliser un entrepôt de données pour procéder à des analyses documentaires lors de l'expérimentation, nous avons eu l'idée d'enrichir l'ENT en proposant un service qui pourrait venir s'intégrer au système d'information et ainsi enrichir la palette des services proposés pour répondre aux besoins des acteurs enseignants-chercheurs. Rappelons que pour favoriser l'intégration de ces services au niveau du système d'information les technologies employées pour le développement des applications reposent sur des logiciels uPortal<sup>4</sup>. Consciente de cette pluralité de possibilités de développement l'Agence de mutualisation des universités (AMUE) travaille à un rapprochement des consortiums dans le souci de pérenniser les développements, d'en favoriser leur réutilisabilité et leur interopérabilité. L'AMUE met l'accent sur l'intérêt de développer des applications autour d'une architecture SOA<sup>5</sup> en faveur des Web services dans l'esprit du web social révélé par l'Internet 2e génération (Mesguich, 2006). Les architectures orientées services permettent d'expliquer la nouvelle complexité de la conception d'applications distribuées. Interopérabilité, standardisation, démarches de conception plus collaboratives et orientées processus, applications composites, solutions de management des processus sont des éléments qui contribuent à modifier l'industrie informatique. L'informatique se réinvente progressivement en repartant d'une conception documentaire de l'information. XML en forme le socle omniprésent. L'impact de XML sur le paradigme informatique se révèle considérable. Les bases de données XML stockent des documents de manière transactionnelle, tout en gardant la capacité de les extraire grâce à de multiples graphes, à l'instar des bases relationnelles.

---

<sup>2</sup> Powerplay comporte EXPLORER et REPORTER qui permettent la création de rapports et la mise en évidence de résultats pertinents pour l'aide à la décision

<sup>3</sup> Transformer crée des hypercubes à partir d'une base multidimensionnelle

<sup>4</sup> uPortal : Framework open source basé sur Java, XML et XSL servant à créer rapidement des portails dédiés aux campus universitaires. Il est développé sous l'égide de JA-SIG. uPortal n'est pas un logiciel prêt à l'emploi, mais plutôt une bibliothèque de classes Java et de documents XML/XSL permettant de développer le portail.

<sup>5</sup> SOA : Service Oriented Architecture ou Architecture Orientée Services

### 3.2 Analyse bibliométrique multidimensionnelle : phase applicative

Nous portons au travers d'une application toutes les données issues de la phase expérimentale pour relever le défi d'utiliser un logiciel en open source : Openi. Il offre des perspectives innovantes quant au traitement du contenu des informations puisqu'il repose sur des schémas XMLA<sup>6</sup> pour l'analyse des données. Nous appréhendons ce nouveau modèle d'analyse pour notre application où est utilisé un langage de développement autorisant la manipulation de bases de données par requêtes MDX<sup>7</sup> en vue d'analyses. Cet outil décisionnel libre repose sur le moteur OLAP Mondrian<sup>8</sup> destiné à la création et à la publication de rapports. Nous avons élaboré des schémas d'analyses en XMLA dont les requêtes MDX permettent de procéder à des analyses multidimensionnelles via une interface web.

#### a) Présentation du système OLAP Mondrian

Les fonctionnalités attendues d'un système décisionnel sont les rapports statiques, les rapports dynamiques, la navigation multidimensionnelle et les indicateurs synthétiques. Selon le site officiel du projet Mondrian, un système OLAP Mondrian se compose de quatre couches travaillant depuis l'utilisateur final vers le centre des données. Ces couches sont : la couche de présentation, la couche de calcul, la couche d'agrégation et la couche de stockage.

#### b) Explicitation d'un schéma Mondrian

D'après Mondrian, un schéma définit une base de données multidimensionnelle. Il contient un modèle logique constitué de cubes, de hiérarchies, de membres et une projection de ce modèle vers un modèle physique. Le modèle logique est composé de balises utilisées pour écrire les requêtes dans le langage MDX. Le modèle physique est la source des données qui est présentée à travers le modèle logique. C'est en général un schéma en étoile qui est un ensemble de tables dans une base de données relationnelle : une table centrale (table des faits) qui contient les données numériques ayant un intérêt pour les analyses et des colonnes clés étrangères vers les autres tables du modèle. C'est à partir de ces autres tables satellites que seront construites les dimensions.

#### c) Application à notre analyse

La récupération de données concerne les données de l'analyse bibliométrique exposée précédemment autour d'un état de l'art sur l'intelligence économique pour offrir une analyse dans une interface web et à distance. Les données récupérées et analysées aboutissent à des vues métiers par type d'acteur. En effet une authentification via Openi par login et mot de passe permet de proposer différentes vues selon le profil de l'acteur qui se connecte.

Nous disposons d'une base Sql en relation avec un schéma en XMLA. A partir des tableaux Excel pour 2001 à 2005 élaborés lors de la phase expérimentale nous alimentons une base sql appelée «foodmart.sql». L'intérêt réside dans le schéma Mondrian qui permet de définir une base de données multidimensionnelle. C'est un fichier eXtensible Markup Language (XML) nommé FoodMart.xml. Dans ce schéma FoodMart on trouve les dimensions, les mesures et le cube. Un cube est une collection de dimensions et de mesures dans un secteur particulier avec une table des faits associée.

Le schéma «FoodMart.xml» identifie le cube propre à l'analyse bibliométrique sous le nom de «synthese» où l'on trouve entre les balises ouvrantes <Cube name="synthese"> et les balises fermantes </Cube> les dimensions et les mesures de ce cube. Nous y avons défini 7

<sup>6</sup> XMLA : Extensible Markup Language Analysis.

<sup>7</sup> MDX : MultiDimensional eXpression.

<sup>8</sup> Mondrian : Serveur OLAP écrit en Java.

mesures agrégées autour des items de colonnes "id\_themes", "id\_evenements", "id\_secteur", "id\_objectifs", "id\_date", "id\_organisateur\_type" et "id\_organisateur\_specialite". Elles sont composées d'une seule hiérarchie, constituée d'un niveau par dimension. Chacun de ces niveaux fait intervenir une colonne de la table\_fait\_synthese.

```
<Cube name="synthese">
  <Table name="table_fait_synt" />
  <Dimension name="evenement" foreignKey="id_fait_synt">
    <Hierarchy hasAll="true" primaryKey="id_fait_synt">
      <Table name="synthese" />
      <Level name="level2" column="evenements" uniqueMembers="true" />
    </Hierarchy>
  </Dimension>
  <Dimension name="theme" foreignKey="id_fait_synt">
    <Hierarchy hasAll="true" primaryKey="id_fait_synt">
      <Table name="synthese" />
      <Level name="level1" column="themes" uniqueMembers="true" />
    </Hierarchy>
  </Dimension>
  <Dimension name="secteur" foreignKey="id_fait_synt">
    <Hierarchy hasAll="true" primaryKey="id_fait_synt">
      <Table name="synthese" />
      <Level name="level3" column="secteur" uniqueMembers="true" />
    </Hierarchy>
  </Dimension>
  <Dimension name="objectif" foreignKey="id_fait_synt">
    <Hierarchy hasAll="true" primaryKey="id_fait_synt">
      <Table name="synthese" />
      <Level name="level4" column="objectifs" uniqueMembers="true" />
    </Hierarchy>
  </Dimension>
  <Dimension name="date" foreignKey="id_fait_synt">
    <Hierarchy hasAll="true" primaryKey="id_fait_synt">
      <Table name="synthese" />
      <Level name="level5" column="date" uniqueMembers="true" />
    </Hierarchy>
  </Dimension>
  <Dimension name="organisateur_type" foreignKey="id_fait_synt">
    <Hierarchy hasAll="true" primaryKey="id_fait_synt">
      <Table name="synthese" />
      <Level name="level6" column="organisateur_type" uniqueMembers="true" />
    </Hierarchy>
  </Dimension>
  <Dimension name="organisateur_specialite" foreignKey="id_fait_synt">
    <Hierarchy hasAll="true" primaryKey="id_fait_synt">
      <Table name="synthese" />
      <Level name="level7" column="organisateur_specialite" uniqueMembers="true" />
    </Hierarchy>
  </Dimension>
  <Measure name="Mesure 1" column="id_themes" aggregator="sum" formatString="#" />
  <Measure name="Mesure 2" column="id_evenements" aggregator="sum" formatString="#" />
  <Measure name="Mesure 3" column="id_secteur" aggregator="sum" formatString="#" />
  <Measure name="Mesure 4" column="id_objectifs" aggregator="sum" formatString="#" />
  <Measure name="Mesure 5" column="id_date" aggregator="sum" formatString="#" />
  <Measure name="Mesure 6" column="id_organisateur_type" aggregator="sum" formatString="#" />
  <Measure name="Mesure 7" column="id_organisateur_specialite" aggregator="sum" formatString="#" />
</Cube>
```

Figure 3 : Schéma xmla du cube «synthese» : table des faits, dimensions, mesures

Une fois les dimensions créées, on liste les mesures de celui-ci. Une mesure<sup>9</sup> est une quantité qu'il est intéressant de quantifier dans ce cube au travers de ses dimensions.

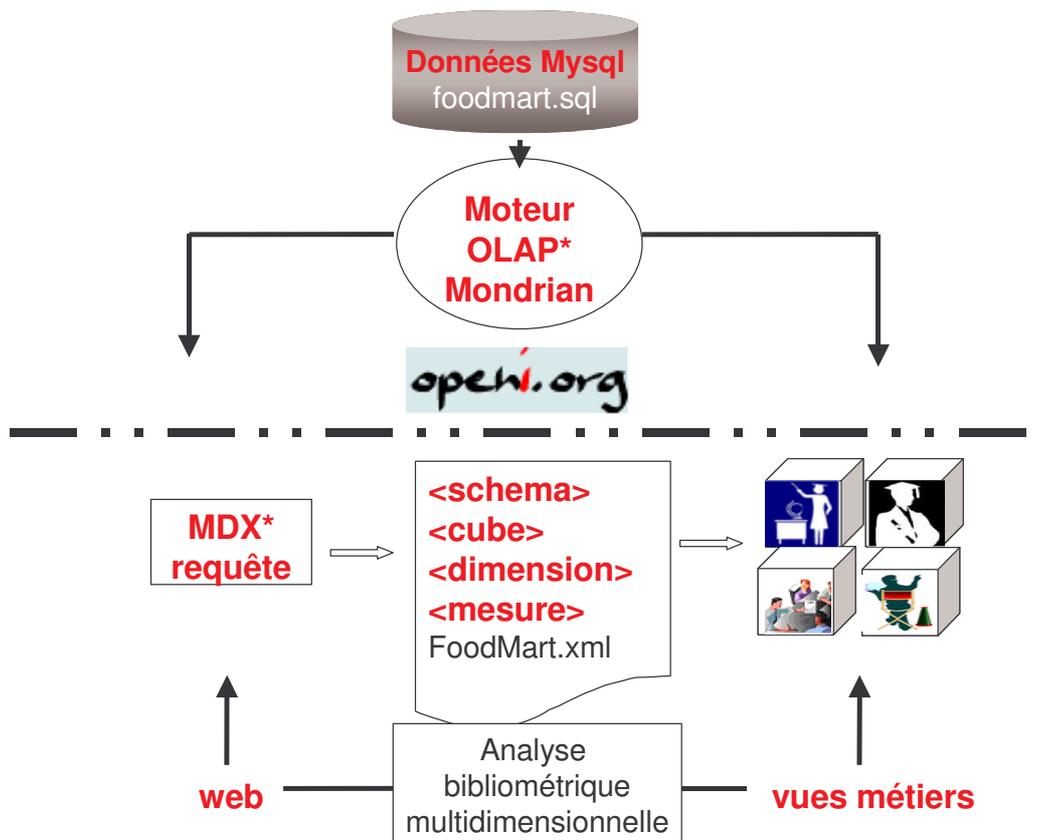
Openi offre un éditeur MDX avec une génération automatique de code lors de la création des analyses. Il offre une gestion des projets. Si l'utilisation interactive des cubes est intéressante, leur structure multidimensionnelle associée à la puissance de ce langage en fait des outils de prédilection pour le reporting opérationnel d'entreprise ou d'administration. Au-

<sup>9</sup> Chaque mesure a un nom, une colonne dans la table des faits et un agrégateur. Cet agrégateur peut être une somme, un maximum ou encore une moyenne.

delà de cet aspect, le MDX permet d'explorer les données des cubes et crée la vraie valeur autour des données et c'est ici que se trouve la plus value.

A partir du cube «synthese» nous faisons une recherche sur les colonnes de l'item «evenement» qui se trouve dans la dimension «Measure». Sur les lignes nous faisons une recherche dans la colonne «evenements» pour effectuer des mesures. En nous promenant sur le graphique, il est possible d'utiliser les fonctions Drill Down<sup>10</sup> et Drill up<sup>11</sup>. Ce procédé de navigation permet la fouille de données pour mettre en évidence que les organisateurs types Prospectives en 2001 et ADBS en 2005 ont consacré le nombre le plus important de journées consacrées à l'événement Intelligence Economique.

Ci-dessous voici schématisé le scénario de l'application développé sous Openi :



**Figure 4 : scénario de l'application sous Openi**

Le recours à une expérimentation puis à une application pour aborder un vocable relatif aux sciences de l'information et de la communication sous l'angle de l'épistémologie témoigne de la biodiversité virtuelle de l'information (Mc Dowell, 2002). Les technologies élaborées que nous employons (entrepôt de données en temps 1, accessible via l'Internet en temps 2) favorisent une lecture des praxis des différentes communautés des sciences de l'information et de la communication.

<sup>10</sup> Drill Down : Forer vers le bas. Aller du général au particulier dans une recherche d'information dans une base de données multidimensionnelle. Détailler selon une dimension, par exemple année, Mois et Semaine.

<sup>11</sup> Drill up : Analyse de données à un attribut parent. Remonter dans la hiérarchie d'une dimension.

#### 4. Conclusion

Nous proposons un modèle original de mise en relation des données numériques et des acteurs qui consiste à agréger un profil à une ressource numérique. Dans notre conception du système d'information stratégique reposant sur un entrepôt de données la description du document électronique devient si fine que le document lui-même devient un «acteur» du système : par une constellation de faits nous faisons converger nos données avec la prise en compte du contexte des utilisateurs. Nous avons montré par ce papier qu'à côté de données chiffrées nous proposons des analyses de contenus documentaires. Pour cela la phase expérimentale a mis en relief la nécessité de travailler autour des métadonnées qui constituent les en-têtes des lignes et des colonnes que l'on retrouve dans le schéma XMLA en vue d'analyses OLAP. En exposant les métadonnées de documents électroniques, nous pouvons passer à une dimension supérieure, c'est à dire à la possibilité de corrélérer des données via des cubes virtuels provenant de systèmes d'information différents (Open Archive Initiative, Système d'Information Documentaire, fichiers de log...) de façon à élaborer un système d'information stratégique orienté vers la prise de décision au service de l'évaluation d'un établissement de recherche par exemple. Notre plate forme <http://perso17.bu.sciences.uhp-nancy.fr:8080/openi/> permet de tester l'analyse bibliométrique multidimensionnelle présentée ici, ainsi que des analyses autour de fichiers de log des services de l'ENT selon des analyses précalculées. Pour mettre en corrélation nos bases de connaissances, nous avons repéré des étapes automatisables en guise de perspectives dans la continuité de nos travaux de recherche.

#### 5. Bibliographie

- Agence de mutualisation des universités, <http://www.amue.fr/Amue/Default.asp>. [Consulté le 30 mars 2007].
- Argyris C., Schon D, 1978. Organizational learning : a theory of action perspective, Adison Wesley, 1978.
- Academic Ranking of World Universities 2006 <http://ed.sjtu.edu.cn/ranking2006.htm> [Consulté le 1er avril 2007].
- Cansell, P. 2003. Actions et méthodes d'IE à Giat Industries : acteurs et outils d'une dynamique d'IE, *IERA*.
- Dalbin S, 2003. La modélisation : pourquoi l'intégrer dans les systèmes d'information documentaire ? *Documentaliste - Sciences de l'information*, 2003, vol. 40, n° 3, p. 226-231.
- Le décisionnel, clé des données structurées : les moteurs de recherche misent sur la capacité de restitution des outils de business intelligence pour remonter les données issues du monde structuré, *OI Informatique*, 2006, juin, p 43.
- Duveau-Patureau V. Le Nouvel enseignant-chercheur : un pédagogue créatif autour de son expertise [http://www.formasup.education.fr/fichier\\_statique/campus/salon/VDPcompetenseigner.ppt](http://www.formasup.education.fr/fichier_statique/campus/salon/VDPcompetenseigner.ppt). [Consulté le 24 février 2003].
- Esup portail : Environnement numérique de travail d'accès intégré aux services pour les étudiants et le personnel de l'enseignement supérieur <http://www.esup-portail.org/> [Consulté le 30 mars 2007].
- E. Fernandez-Medina E., Trujillo J., Villarroel R., Piattini M, 2006. Access control and audit model for the multidimensional modeling of data warehouses In *Decision Support Systems* vol 42, 2006, p. 1270–1289.
- Franco J.M. 1997. Le Data Warehouse : objectifs, définitions, architectures . Eyrolles.

- Guilhon A. 2003. Le processus d'intelligence économique et l'identité de la PME, *Economica*.
- Guilhon B., Levet J.-L, 2003. De l'intelligence économique à l'économie de la connaissance. *Economica*.
- Lainé-Cruzel S. Appropriation, mutualisation, expérimentations des technologies de l'information scientifique et technique. Paru dans : Partie 1 <http://ametist.inist.fr/personne.php?id=151&type=auteur>. [Consulté le 30 mars 2007].
- Mc Dowell Ashley. 2002. Trust and information: the role of trust in the social epistemology of information science= Confiance et information: le rôle de la confiance dans l'épistémologie sociale de la science de l'information, In *Social epistemology*, vol. 16, no 1, pp. 51-63.
- Mesguich V. Le Web 2.0 démystifié : principes, définitions et atouts pour les professionnels. *Netsources*, 2006, n° 64.
- Oubrich M., 2003. Processus d'intelligence économique : transformer l'information en connaissance, *AIM 2003*, Grenoble.
- Peguiro F. , 2001. Accès à l'information sur Internet, pratiques et tendances des utilisateurs : dans un contexte de documentation électronique, Université de Nancy-Metz.
- Peguiro F. , 2006. Application de l'Intelligence Economique dans un Système d'Information Strategique universitaire : les apports de la modélisation des acteurs. Thèse : informatique : Université Nancy II .
- Peguiro F.,Thiery O., 2005. Modélisation des acteurs et des ressources : application au contexte d'un SIS universitaire, *ISKO2005*, Nancy.
- Peguiro F., Thiery O., 2005. Système d'information stratégique dédié à l'environnement universitaire, *COSI2005*, Bejaia.
- Prensky, M. 2004. Capturing the Value of "Generation Tech" Employees, <http://www.strategy-business.com/enewsarticle/enews063004>, [Consulté le 12 avril 2006].
- Rongeat D. , 2007. Intégration dans les ENT, *Esup Days 26 janvier 2007*.
- Rosnay Joël de, 2006. La Révolte du pronétariat, Fayard, 2006.
- Schéma directeur des espaces numériques de travail, Ministère de la jeunesse, de l'éducation nationale, et de la recherche, 2004, <http://www.educnet.education.fr/chrgt/SDET-v1.doc>.
- Sweeny, R. Creating WOW ! Services for millenials », <http://www.library.njit.edu/staff-folders/sweeney/Millennials%203-26%202004%20Test/Millennials%20Web%20Site.ppt>, [Consulté le 12 avril 2005].
- Thivant E. et Bouzidi L. 2005. Les pratiques d'accès à l'information : le cas des concepteurs de produits de placements financiers, *Revue électronique suisse de science de l'information*, n°2, p.7-34.
- Varandat M, 2005. Avez-vous nommé votre gouverneur de données ? *OI Informatique*, 2005, octobre, p. 44-46.