HAL open science

Un SIS pour les universités – Modélisation de l'acteur et méta données

Frédérique Peguiron

▶ To cite this version:

Frédérique Peguiron. Un SIS pour les universités – Modélisation de l'acteur et méta données. 2005. inria-00000355

HAL Id: inria-00000355
https://inria.hal.science/inria-00000355v1

Submitted on 27 Sep 2005

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.

Un SIS pour les universités

Modélisation de l'acteur et méta données

Frédérique Peguiron

LORIA
B.P. 239
54506 Vandoeuvre-les-Nancy
Frederique.Peguiron@loria.fr

RÉSUMÉ *Une université abrite différents acteurs qui ont recours à des systèmes de ressources documentaires, des systèmes de production d'information, des systèmes de recherche d'information. Le recours à l'entrepôt de données(ED) permet de proposer des solutions pour faire évoluer un Système d'Information(SI) en un Système d'Information Stratégique (SIS). La prise en compte de notre modèle de représentation de l'utilisateur RU=(T, F, B, A) est traduit par des méta données. Nous élaborons également les méta données propres à l'(ED) qui définissent les méta données structurelles et d'accessibilité propres au système de pilotage. Afin de procéder au mieux au développement de notre contribution au (SIS), la méta modélisation de l'(ED) permet d'élaborer un schéma directeur pour la construction de l'(ED).*

ABSTRACT. *A university have various actors who use systems of documentary resources, systems of production of information, systems of information search.. The use of a datawarehouse (DW) makes it possible to propose solutions to move a System of Information (SI) in a Strategic Information System (SIS). Our user modeling RU=(T, F, B, A) is translated by meta data. We also work out the meta data specific to the (DW) which define the meta data structural and of accessibility of the system of control. In order to carry out as well as possible the development of our contribution to (SIS), the meta modeling of the (DW) makes it possible to prepare a directing diagram for the construction of the (DW).*

MOTS-CLÉS : *Université, système d'information stratégique, classification des utilisateurs, modélisation de l'utilisateur, entrepôt de données, méta données, méta modélisation.*

KEY WORDS : *University, strategic information system, user classification, user modeling, data warehouse, , meta data, meta modelling.*

1. Introduction

Une université abrite différents acteurs qui ont recours à des systèmes de ressources documentaires, des systèmes de production d'information, des systèmes de recherche d'information. Au sein de cette même université, cohabitent de nombreux Systèmes d'Information (SI) spécifiques aux besoins des composantes qui la constituent. Ces (SI) éparses abritent des informations qui peuvent être utiles aux composantes voisines. Le recours à l'entrepôt de données permet de proposer des solutions pour faire évoluer un Système d'Information en un Système d'Information Stratégique (SIS) voire en un Système d'Information Décisionnel (SID). Le transfert d'un (SI) en (SID) repose sur les bases métiers orientées vers les acteurs. Cela suppose une réflexion sur deux problèmes : celui des utilisateurs et des méta données. Nous articulerons nos propos autour de deux parties. En premier lieu, nous aborderons la modélisation des acteurs. Après quoi, nous aborderons les méta données et la méta modélisation.

Aujourd'hui, les technologies décisionnelles actuelles permettent aux dirigeants de bénéficier d'informations souples sur son environnement. Cependant, un effort important doit être apporté à la qualité des données qui sont désormais accessibles à un plus grand nombre de personnes, de même qu'à la qualification des données afin de permettre aux dirigeants d'accéder à une information claire. Les Systèmes d'Information Stratégique (SIS) doivent réaliser la synthèse et l'agrégation de données stratégiques issues de divers services. Chacun de ces services et des systèmes informatiques qui leur sont associés ont leurs propres contraintes, leur propre sémantique. Un Système d'Information Décisionnel (SID) est un système qui permet aux décideurs d'une institution de disposer d'informations pertinentes et d'outils d'analyse puissants pour les aider à prendre les bonnes décisions au bon moment. Le suivi d'une politique d'établissement peut se faire par le partage de tableaux de bord et autres indicateurs. D'après Inmon- William H.[**17**] «Le Data Warehouse est une collection de données orientées sujet, intégrées, non volatiles et historisées, organisées pour le support d'un processus d'aide à la décision». L'Entrepôt de données est entièrement construit selon une approche dimensionnelle, c'est-à-dire qui fait appel aux techniques qui favorisent l'analyse multidimensionnelle des données. Après ces rapides définitions qui rappellent que l'entrepôt de données ou data warehouse est une base de données d'aide à la décision qui est maintenue séparément de la base opérationnelle de l'organisation, nous allons présenter rapidement les différents projets nationaux. Afin de procéder au mieux au développement de notre contribution au Système d'information, il est nécessaire d'élaborer un schéma directeur pour la construction de l'entrepôt de données, ainsi que la construction de son méta-dictionnaire. Ce processus passe par l'adoption d'une méthodologie pour la construction d'un référentiel de métadonnées. Les métadonnées sont des informations sur les données indispensables à une agrégation efficace d'un entrepôt de données, un ensemble d'informations d'administration et de suivi pour le projet décisionnel. Plus précisément, les métadonnées représentent toutes les informations nécessaires à

l'accès, à la compréhension de l'exploitation des données : sémantique, origine, règles, agrégation, stockage, format, utilisation.

1.1. L'existant

En analysant "l'Existant", quatre éléments importants ressortent et constituent le cadre de base pour élaborer une stratégie :

Agence de mutualisation des universités

Sur un plan national l'agence de mutualisation[**3**] des universités mène une étude pour produire des cahiers des charges autour de ses futurs produits pour l'année 2006, plus spécifiquement autour de la Scolarité et de la Gestion en vue de l'élaboration fonctionnelle, organisationnelle et technique du SI ou Système d'information de gestion d'un établissement.

L'entrepôt de données

Au sein de l'agence de mutualisation des universités une équipe sous la conduite de Sibylle Rochas, élabore un entrepôt de données à partir d'Harpège, Apogée et Nabuco afin d'aider au pilotage des universités. Ce projet[**12**][**25**] aboutit par la mise à disposition d'un extracteur de données et d'un méta-dictionnaire pour les sites pilotes.

L'audit

Le cabinet de consultant Cap Gemini Ernst et Young[**6**] mène un audit sur les Etudes préalables à l'élaboration d'un système d'information et de gestion afin d'optimiser les performances d'un établissement universitaire. Nous retrouvons dans la Tranche 1 de cette étude les concepts propres à l'entrepôt de données qui doivent nous guider dans l'élaboration de notre démarche qui sont notamment : paramétrer le système de pilotage, définir les objectifs, ajuster les tableaux de bord, alerter...

Esup

Esup portail[**14**] est la constitution d'un consortium d'universités auquel adhèrent d'ores et déjà : l'Université de Valenciennes, l'Université Nancy 2, l'Université Henri Poincaré Nancy 1, l'Université de Rennes 1, l'Université Paul Sabatier Toulouse 3 bientôt enrichi d'autres universités : Université du Littoral, Université de Versailles, INPL, INPT, Université du Havre, Université Antonine (AUF), Université Rennes 2, IUFM de Bretagne, Université de Franche-Comté, Centre Universitaire de Formation et Recherche Jean-François Champollion (ALBI)

Rappelons qu'un espace numérique de travail désigne un dispositif global fournissant à un usager un point d'accès à travers les réseaux à l'ensemble des ressources et des services numériques en rapport avec son activité. Il est un point d'entrée pour accéder au système d'information de l'établissement. L'établissement d'enseignement est le périmètre de référence de l'espace numérique de travail du point de vue de l'utilisateur. L'espace numérique de travail s'adresse ainsi à l'ensemble des usagers, étudiants, enseignants, personnels administratifs et techniques. Cela ne signifie pas que les services et ressources sont exclusivement fournis par l'établissement: l'espace numérique de travail doit au contraire favoriser leur mutualisation, au niveau inter établissements, avec les partenaires publics et privés, en France, en Europe ou au niveau international. Esup portail contient les

briques assurant l'accès aux applications métiers mises en place dans les universités (comptabilité, scolarité, GRH, enseignement à distance, gestion documentaire...). Il s'appuie sur le système d'information de l'établissement qui doit être conçu en cohérence avec les applications de gestion propres à l'établissement. Esup Portail retenu par l'université Nancy 2 implique d'avoir un système d'information complet, cohérent et structuré. Cela suppose une base d'utilisateurs avec droits associés, ainsi que la définition de profils types et de leur environnement standard. Parallèlement à l'analyse des besoins spécifiques, il faut mettre en place une information et formation des utilisateurs. L'architecture technique d'Esup portail est conçue sur un principe de construction de briques. Cette fabrication nous assure de pouvoir nous intégrer à l'existant. La figure 1 permet de situer la n^{ème} brique au SI, emplacement où nous pouvons nous intégrer.

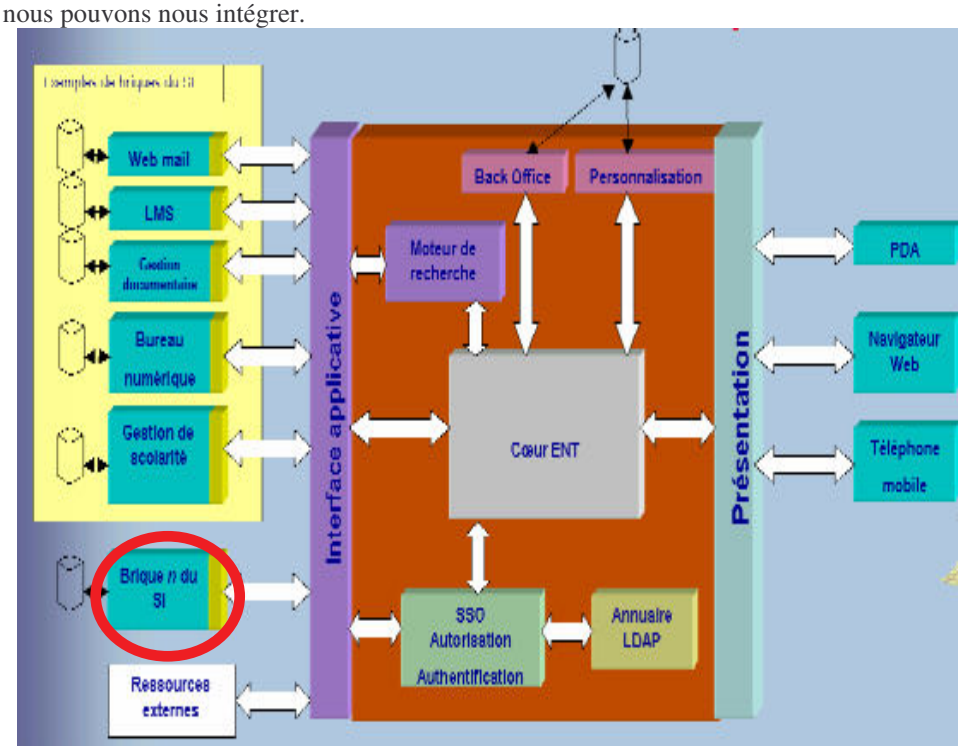

Figure 1. Place de notre brique au sein du SI de l'université

L'OLAP ou Online Analytical Processing est une technique informatique d'analyse multidimensionnelle, qui permet aux décideurs, d'avoir accès rapidement et de manière interactive à une information pertinente présentée sous des angles divers et multiples, selon leurs besoins particuliers. A titre d'exemple on peut représenter de façon graphique des informations contenues dans une base de

données, sous la forme d'un cube à plusieurs dimensions, lequel cube permet d'analyser ces données sous différents angles, grâce à l'organisation de celles-ci en axes d'analyses et en variables à analyser.

Un entrepôt de données donne naissance, par filtrage non plus par rapport aux dimensions mais par rapport à des profils utilisateurs, à des bases métiers (ou data marts). Ce sont des sous bases de l'entrepôt de données destinées à une fonction de l'organisation : marketing, financier... Elles sont alimentées périodiquement, reposent sur une vue multidimensionnelle des données, et elles sont non modifiables par les utilisateurs. La modélisation de l'utilisateur intervient dans deux contextes de S-IS : pour la constitution des informations pertinentes et pour l'exploitation du S-IS. Dans le premier contexte, il s'agit d'utiliser le modèle de l'utilisateur comme un profil pour le filtrage d'information à intégrer dans le S-IS. Dans le deuxième contexte, il s'agit d'utiliser le modèle comme outil d'adaptation du comportement du système aux comportements de l'utilisateur. Pour illustrer notre brique[**28**] au Système d'Information, la figure suivante permet de corréler les besoins des utilisateurs et les bases métiers.

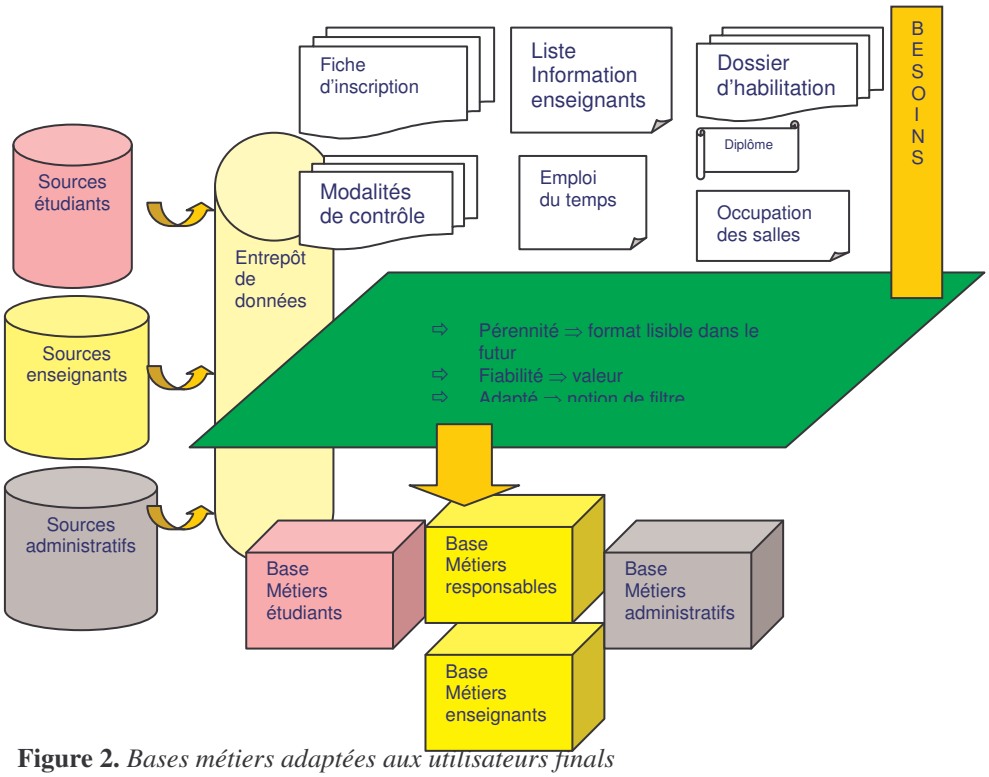

Figure 2. *Bases métiers adaptées aux utilisateurs finals*

Notre approche de filtrage de l'information par le profil de l'utilisateur[**9**] est donc fondée sur la modélisation de l'utilisateur en intégrant des attributs d'identification de l'utilisateur, de son comportement et du contexte d'utilisation des

informations qu'il cherche. Le processus du filtrage de l'information lors de la constitution de l'entrepôt de données s'opère par un filtrage thématique et dans la constitution des bases métiers par un filtrage fonctionnel.

2. Modélisation de l'acteur

2.1. *Contexte*

Nous avons pu voir au cours de nos recherches que les logiciels administratifs au sein des universités, ont des limites au niveau du pilotage. Un système d'information informatisé permet une exploitation de données sans ambiguïté et exclut l'aléatoire dans les processus et les événements. Apogée[1], par exemple, qui permet l'inscription administrative et pédagogique de chaque étudiant, requiert une modélisation dépendante de chaque établissement. Apogée permet d'avoir la répartition des étudiants par diplôme en croisant les instances de l'objet «individu» et les instances de l'objet «diplôme». Toutefois, la modélisation étant très figée dans Apogée, des requêtes de pilotage pertinentes ne peuvent pas être définies dynamiquement au fur et à mesure des besoins. Comme par exemple l'impossibilité de répondre à la question : le nombre de redoublants pour le 2e cycle, car le système n'a pas pris en compte la notion de redoublant et ne peut pas évoluer pour l'appréhender. Néanmoins, lors de la mise en route du processus de modélisation des acteurs : étudiants, enseignants, administratifs nous avons mis en évidence que nous pouvons récupérer des données au niveau de ce logiciel afin de faire le lien entre enseignant, enseigné et enseignement. Nous évoluons dans un contexte universitaire où il s'agit de mettre en relation information et acteurs du SI–SIS. Nous allons voir comment la modélisation des acteurs permet d'améliorer les réponses.

2.2. *Modélisation des acteurs*

Nous dressons une classification[**29**] des acteurs[**30**] où nous disons qu'un utilisateur (**U**) est représenté par un type (**T**), besoins (**B**), fonctions (**F**), et activités (**A**) selon ce modèle :

$$\boxed{\mathbf{RU = (T, B, F, A)}}$$

Nous allons développer ce modèle de représentation d'un utilisateur. Notre étude tient compte de l'espace numérique de travail (Esup portail) qui se met en place à l'Université de Nancy 2. Le schéma directeur des espaces numériques de travail précise les usagers amenés à intervenir dans cet environnement. Ceci permet de mentionner les utilisateurs concernés par notre réflexion. Nous nous appuyons dans ce travail, sur les idées développées par des pédagogues reconnus et aussi sur

[1]Apogée, Application pour l'organisation et la gestion des étudiants, apporte des réponses précises en matière de clarification de l'offre de formation, d'amélioration de l'accueil des étudiants, de gestion de la scolarité et de pilotage de l'établissement.

l'apport des TIC (Technologies de l'Information et de la Communication). L'ingénierie pédagogique, dont le rôle est la transmission, est en mesure de passer à une ingénierie d'apprentissage en favorisant la création. Les apprenants peuvent passer de produits à consommer à des créations de services. La prise en compte de notre modèle de représentation de l'utilisateur RU=(T, F, B, A) se situe en amont du processus de schéma directeur pour l'élaboration de l'entrepôt de données. Dans les paragraphes suivants, nous allons développer les items T : type, F : fonctions, B : besoins et A : activités des acteurs par rapport au (SI).

2.2.1. *Type*

L'item **T** représente le type d'acteur pour l'enseignement supérieur que nous résumons de la façon suivante : **T** {Etudiants, Chercheurs, Enseignants, Responsables, Personnels, Partenaires, Administrateurs}. Nous utilisons le formalisme UML pour modéliser les types d'acteurs. Empruntons à UML les diagrammes de classe et d'objets pour nous permettre de recenser des objets et des classes dans les données fournies sur les maquettes de cours tirées d'Apogée. L'étape de modélisation de classes d'objets permet de faire apparaître des attributs et des valeurs selon le schéma suivant :

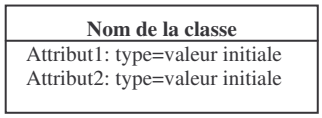

Figure 3. *Formalisme de représentation*

La modélisation des types d'acteurs sur la figure 4, permet de dresser des catégories d'acteurs et des sous-catégories d'acteurs. Cela permet d'introduire les notions de groupes, de sous-groupes et de leurs rôles respectifs, notions qui seront développées ultérieurement.

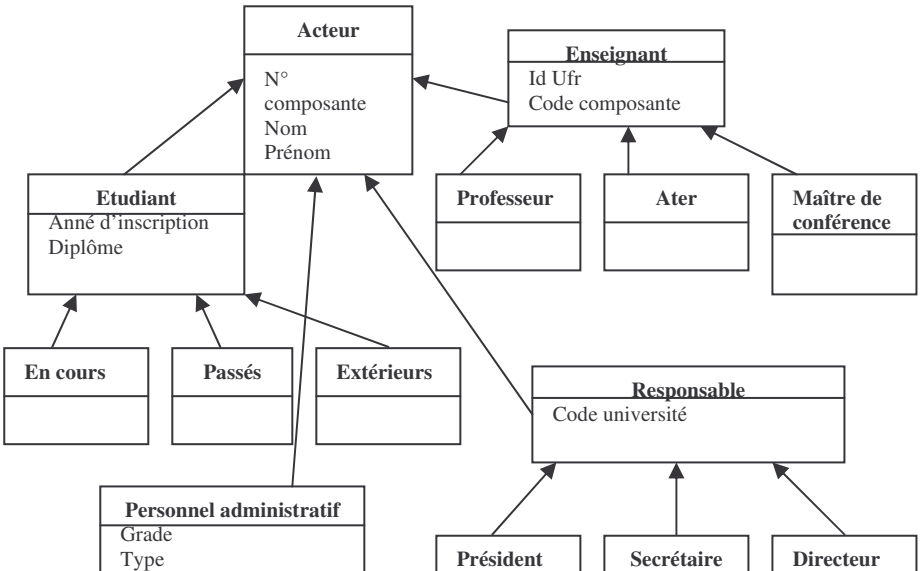

Figure 4. *Modélisation des types d'acteurs*

Concentrons nous sur l'acteur «étudiants» de la figure 4. D'après[**21**] le manuel d'utilisation de l'infocentre pilotage pour Apogée, l'univers étudiant est représenté par les classes suivantes : **individu**, inscription pedagog. etape, inscription admin. annuelle, inscription admin. etape resultats aux elements, resultats aux etapes, **resultats aux diplomes**, stages, troisieme cycle, couverture sociale, cursus pre-universitaire, **diplome**, diplome autre cursus, **cursus**, blocage, indicateurs. Chaque classe comporte des objets. Rappelons que l'un de nos objectifs consiste à améliorer l'information proposée à l'utilisateur. Nos trois classes d'objets sélectionnées : **diplomes, individu**, **cursus** doivent nous permettre, par exemple d'aider l'acteur étudiant à améliorer ses connaissances autour d'une discipline.

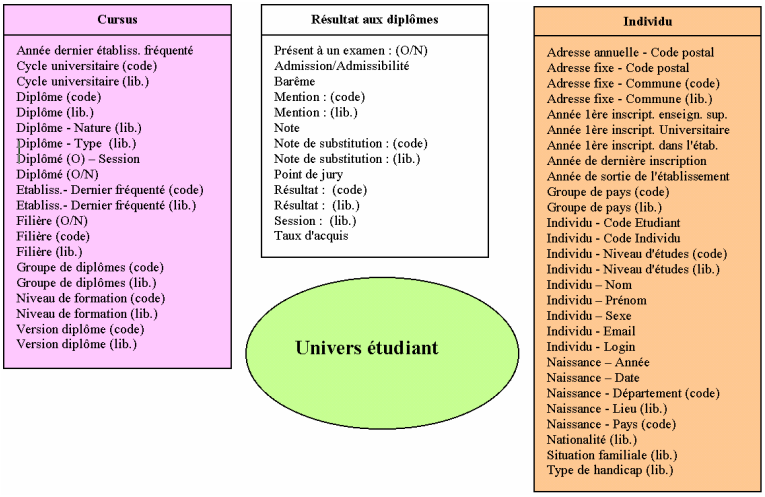

Figure 5. *Exemple de 3 classes d'objet modélisant l'acteur étudiant*

Pour l'acteur étudiant, choisissons la classe **individu**, pour voir comment l'on peut tirer parti des données existant dans les champs ou dans les vues partielles, en vue de les réutiliser pour lui permettre d'améliorer ses recherches.

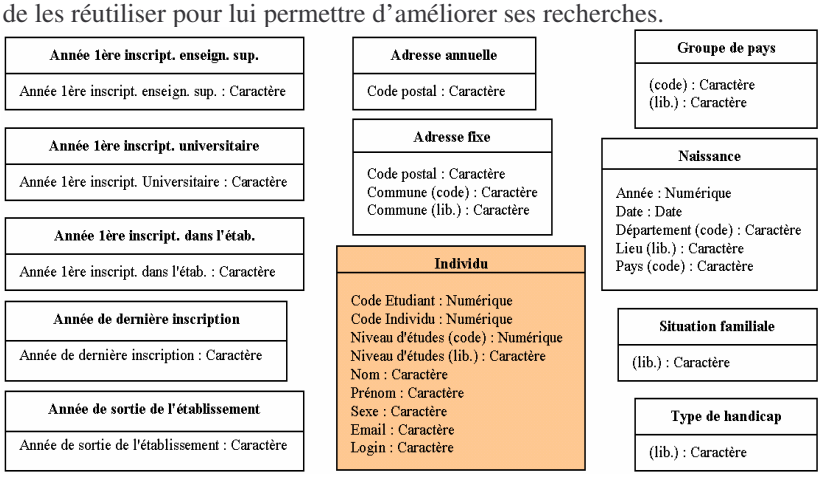

Figure 6. *Développement de la classe individu pour l'acteur étudiant*

La figure 6 fait apparaître des données propres à l'identification de l'étudiant, sa localisation géographique, son niveau d'étude, son login, son historisation d'inscription dans l'établissement. Avant de poursuivre sur la modélisation d'une

composante, présentons et développons la classe **diplôme** formée d'U.E. (unités d'enseignement) :

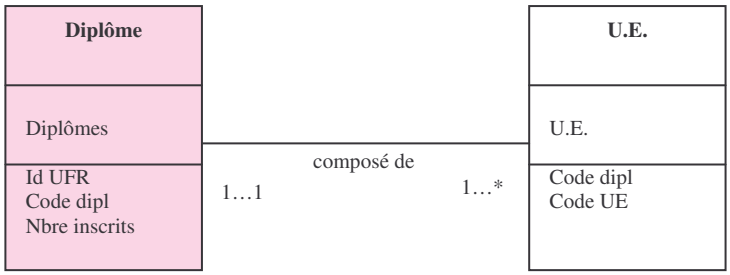

Figure 7. *Classe diplôme*

En développant la classe cursus de la figure 5, les objets autour du diplôme nous renseignent sur les modules. Auparavant, analysons une composante pour observer les diplômes proposés. Voici les diplômes proposés par la composante UFR Mathématiques Informatique de Nancy 2 pour l'année 2002.

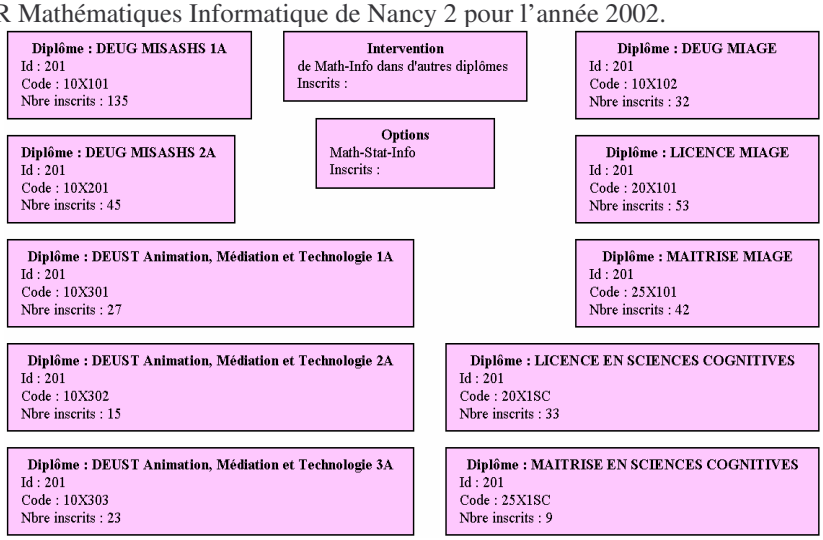

Figure 8. *Modélisation des interventions et diplômes de la composante UFR MI de Nancy 2*

Sur le même principe, choisissons le diplôme DEUG MISASHS Info 1A de la composante UFR MI de Nancy 2, pour développer le contenu des U.E.

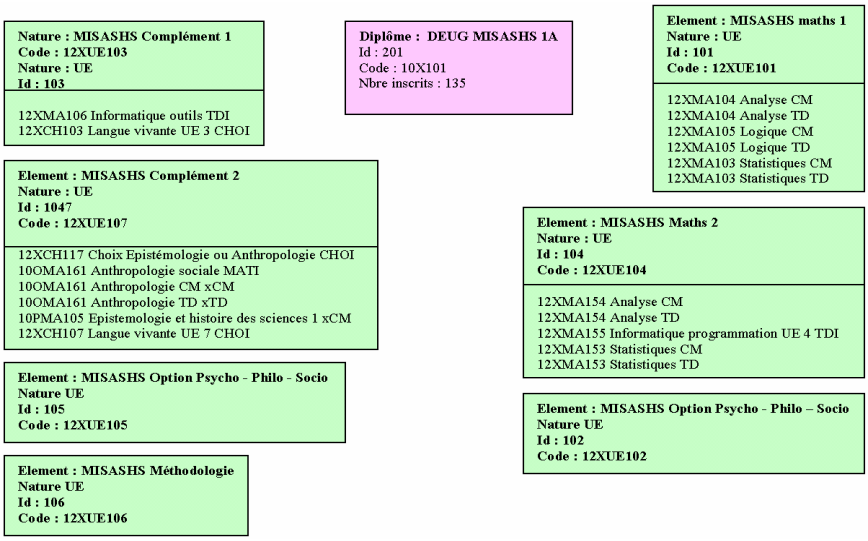

Figure 9. *Modélisation d'un diplôme DEUG MISASHS 1A*

Récapitulatif :

La modélisation des types d'acteurs, nous a permis de modéliser l'acteur étudiant, de développer les classes individu et diplôme. Les attributs et les valeurs nous donnent des données chiffrées comme, par exemple, le nombre d'étudiants pour une discipline ; ils nous permettent également de mettre en valeur des notions en procédant par des calculs par association : comme cerner les compétences d'un étudiant autour d'un domaine ou de calculer le nombre de redoublants.

Domaine d'application :

Notre domaine d'application [**30**] met en situation un enseignant qui doit concevoir une même formation à deux publics d'étudiants appartenant à des UFR différentes et un étudiant amené à améliorer ses compétences autour de modules. L'utilisateur, dont le profil est décrit dans des méta-données, obtient des informations personnalisées. Ces informations peuvent concerner des ressources documentaires ou des informations économiques et financières. En fonction de son rôle, l'utilisateur pourra visualiser des niveaux d'information. Cet exemple met en évidence que les étudiants et les enseignants ont des besoins, des fonctions et des activités spécifiques[**29**].

2.2.2. *Besoins*

L'enseignant[**28**] est placé à la fois dans un environnement d'innovation et aussi dans un environnement réglementé. Dans un contexte universitaire, l'enseignant a besoin de répondre aux besoins des étudiants, de préparer des formations, de créer

des formations, de connaître l'impact d'une formation, d'évaluer le contenu des formations, d'adapter les formations, d'évaluer les connaissances des étudiants. Parallèlement, l'enseignant cherche à rationaliser la préparation des cours, de voir garantie la propriété intellectuelle et d'obtenir une reconnaissance de ses cours au niveau de sa carrière. On peut renseigner un certain nombre d'éléments de l'item besoins pour les acteurs étudiant et enseignant. Cet item peut s'enrichir au cours du temps : $\mathbf{B}_{\text{étudiant}}$ {s'inscrire, s'exercer, se former, rechercher emploi, rechercher stage} ; $\mathbf{B}_{\text{enseignant}}$ {exercer, former, corriger, recenser, évaluer, budgétiser, déployer, planifier, se conformer aux textes officiels}.

2.2.3. *Fonctions*

Parallèlement aux besoins, les enseignants et les étudiants ont des fonctions qui leur confèrent un rôle. L'enseignant est amené à distribuer des devoirs, décrire des tâches à effectuer, évaluer, réceptionner des travaux. Enseignants et étudiants peuvent être amenés à utiliser des ressources : pour créer ou publier des pages web, pour visionner des documents vidéo, pour prendre connaissance des objectifs visés de l'auteur, pour lister des articles, des URL, des références de documents en rapport avec le thème traité, pour historiser l'état d'avancement d'un travail. Par les fonctions de communication, les acteurs peuvent : recourir à un calendrier, discuter, utiliser un système de messagerie. Lors du processus de création de cours, les enseignants ont des fonctions d'organisation. Ils peuvent être amenés à : mettre à disposition sur le web des ressources pour les étudiants, partager une bibliothèque de documents, gérer et suivre les stages étudiants. Les fonctions de supervision des enseignants permettent : de suivre et encadrer des projets étudiants, d'évaluer les étudiants, de recueillir des travaux d'étudiants sous format numérique, de tutorer à distance des étudiants. Nous pouvons spécifier un certain nombre d'éléments de l'item fonctions ainsi pour les étudiants et les enseignants : $\mathbf{F}_{\text{étudiant}}$ {apprendre, créer, intégrer, vérifier} ; $\mathbf{F}_{\text{enseignant}}$ {créer, enseigner, diriger, missionner, organiser, gérer, conseiller, superviser}. Cette liste n'est pas close.

2.2.4. *Activités*

Les activités des acteurs concernent leurs activités lors de l'utilisation du système d'information. Ils peuvent être amenés à rechercher, télécharger des dossiers, comprimer des fichiers, annoter des images ou du texte, indexer des documents, consulter des notes. Ils doivent parfois disposer d'un certain degré d'interactivité avec le système. Cette phase correspondra plus à l'exploitation des informations. Il peut s'agir d'informations à déposer, d'informations trouvées, d'information à faire connaître à des publics ciblés. A côté de ses tâches d'intendance (gestion de l'effectif, enregistrement des résultats), l'enseignant est amené à participer à : la mise en forme des connaissances, la réalisation d'un glossaire et d'un index, la conception et la réalisation de la navigation, la réalisation de contrôles de connaissance. Ceci met en exergue de nouvelles pratiques quant aux activités des acteurs. Nous résumons les activités des étudiants et des enseignants

dans l'item activités: $\mathbf{A}_{\text{étudiant}}$ {déposer, explorer, interroger, analyser, synthétiser, annoter} ; $\mathbf{A}_{\text{enseignant}}${déposer, indexer, diffuser, explorer, interroger, analyser, synthétiser, annoter}. Ces activités interviennent lors de la création et la consultation des ressources d'information. Le modèle des acteurs nous a permis de faire ressortir des éléments utiles pour la suite de notre article.

3. Métas données et méta modélisation

3.1.1 *Qu'est ce qu'une métadonnée*

Les métas données sont des informations sur les données indispensables à une exploitation efficace d'un data warehouse (DW). Le terme de méta données est couramment défini comme "des données concernant les données". Les méta données décrivent le schéma d'un (DW) et les données individuelles. Dans l'entreposage de données, les méta données sont classées d'après leur objet et le public auquel elles s'adressent. La donnée est forcément liée à d'autres objets du système d'information ; il est donc également nécessaire de représenter, de décrire et de stocker ces interactions avec d'autres données. Notre préoccupation principale est donc la prise en compte de l'acteur dans la construction puis l'exploitation d'entrepôts de données. C'est elle qui permet d'être garant de la cohérence fonctionnelle et technique et de la qualité des informations fournies aux utilisateurs métier par le biais de l'entrepôt de données. Les méta données permettent de décrire l'ensemble des règles, des définitions, des transformations et des processus associés à une donnée. On peut donc, par ce moyen, suivre l'évolution des données depuis le système source jusqu'à leur restitution. On est capable, dès lors, d'avoir une vision unifiée des données, transverse à tous les métiers de l'organiation.
Les principales informations sont destinées :

- A l'utilisateur (informations sur la sémantique des données utilisées et leur localisation dans l'entrepôt de données),
- Aux équipes responsables des processus de transformation de données de l'environnement de production vers l'entrepôt de données (informations sur la localisation de la donnée dans les systèmes de production, sur la description des règles et des processus de transformation),
- Aux équipes responsables des processus de création des données agrégées à partir des données détaillées,
- Aux équipes d'administration de la base de données,
- Aux équipes de production (informations sur les procédures de chargement, historiques des mises à jour, etc..).

3.1.2. *Typologie des méta données*

La description des méta données[39] permet d'éviter la perte de données essentielles. Le processus de description des méta données permet d'inscrire les

méthodes de collecte et de traitement des données. Cela favorise l'utilisation ultérieure des données. Les méta données fonctionnelles[31] liées à une donnée peuvent être :

- Les métas données des systèmes source (structures de données initiales)
- Les métas données SGBD (ex : tables systèmes décrivant le DW)
- Les métas données des outils frontaux (interfaces utilisateurs)
- Les métas données des données du DW

Les méta données techniques permettent de répondre aux questions suivantes :

- Que représente cette donnée ? (sémantique)
- D'où vient la donnée ? Qui l'a créée (origine)
- Comment est-elle calculée ? (règle de calcul)
- Quel est son format ? (stockage)
- Avec quelle fréquence est-elle mise à jour ? (utilisation)
- Qui est responsable de cette donnée ? (administrateur)

La figure 10 contextualise les métas données dans le système de pilotage. On remarque deux types de métas données :

- structurelles : elles décrivent la structure et le contenu du DW (méta schéma)
- accessibilité : elles représentent le lien dynamique entre l'entrepôt et les utilisateurs (référentiel de l'entrepôt ou DW ou Repository)

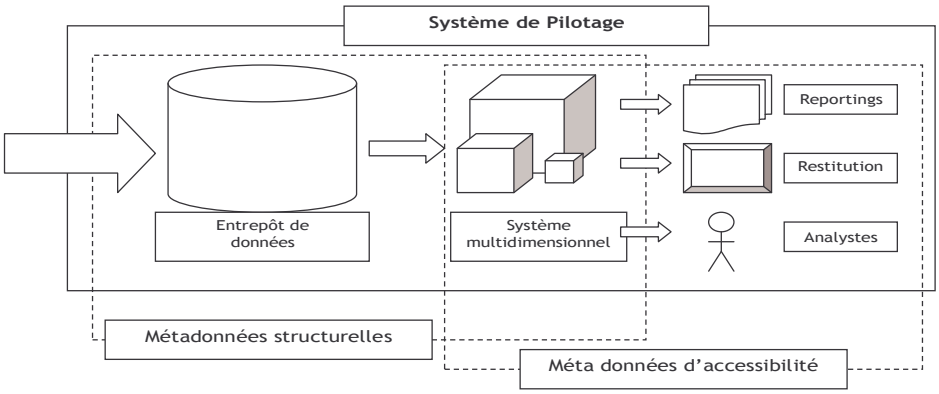

Figure 10. *Métas données structurelles et d'accessibilité*

3.1.3. *Référentiel de l'entrepôt de données*

Il existe plusieurs manières de représenter[40] et de stocker ces méta données. Pour en faciliter l'accès et la manipulation, on pourrait utiliser une fiche contenant une description de la données et de ses liens. Mais généralement, on utilise un symbolisme courant en informatique : les modèles de données au sens Merise du

terme. Les méta données sont exprimées sous forme de modèles conceptuels et de modèles logiques de données. Ainsi, on obtient une vision synthétique permettant de naviguer entre les différentes données. Néanmoins, pour affiner cette vue, on peut utiliser, à l'instar de l'entrepôt de données, des modèles multidimensionnels pour représenter le référentiel (selon axe d'origine, de sémantique, etc.). C'est souvent le lien entre le modèle opérationnel et le modèle décisionnel qui est décrit par les méta données. Or les modèles opérationnels sont dits locaux : dans un entrepôt nous avons besoin d'élargir leur vision afin de consolider une vue au niveau institution. L'idée est d'obtenir ainsi une représentation unique d'une donnée quel que soit le secteur ou l'acteur de l'université. L'ambition avouée : fédérer l'ensemble.

Pour cette raison, le besoin d'une fonction de gestion des méta données apparaît clairement lors de la mise en place d'un projet décisionnel, fonction que l'on va appeler souvent «administration des données». L'ensemble des outils nécessaires à la mise en oeuvre d'une fonction est le référentiel de l'entrepôt de données. Cet outil doit être apte à gérer et à construire un modèle d'institution, enrichi de données de suivi propres à l'entrepôt.

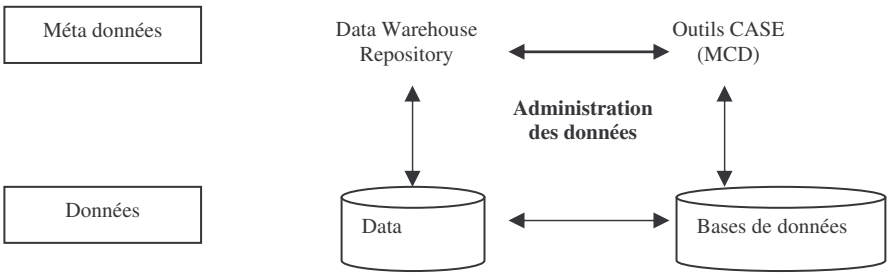

Figure 11. *Vision du référentiel de l'entrepôt de données*

3.2.1. *Métamodélisation[37]*

Les métas connaissances dont nous avons besoin pour obtenir un S-IS adaptable sont des connaissances sur 1) la représentation d'un problème à résoudre, 2) la représentation des comportements de l'utilisateur et 3) la représentation des informations du domaine d'application. Les métas connaissances de 1) et 2) sont liées à la modélisation de l'utilisateur et sont indépendantes des domaines d'application. Ces métas connaissances constituent en fait le méta modèle de l'utilisateur. Le méta modèle de l'utilisateur que nous proposons se présente ainsi : RU = (T, B, F, A).

Pour répondre à notre objectif, nous étudions le SIS sous forme de trois niveaux de modélisation que nous décrivons dans la figure 12.

– Au niveau de la couche 3, on insiste sur le méta modèle de l'utilisateur : en capturant des connaissances sur le décideur pour les mettre dans les métas données de l'entrepôt afin de construire une base métier spécifique à un groupe de décideur ou mieux encore à un décideur particulier.

– Au niveau de la couche 2 : en stockant parmi les métas-données du système, une représentation explicite de la structure des différentes bases métiers, ou plutôt celle des utilisateurs à qui elles sont destinées.
– Au niveau de la couche 1, on prend en compte l'acteur (utilisateur) dans la construction puis l'exploitation de l'entrepôt de données.

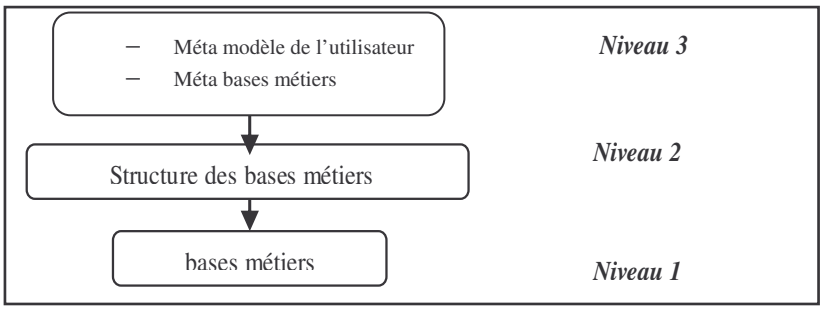

Figure 12. *Les trois niveaux de modélisation d'un S-IS*

Ce méta modèle fait partie du troisième niveau de la hiérarchie de modélisation que nous proposons pour la modélisation d'un S-IS. Il permet de décrire d'une part l'objectif de l'utilisateur, d'autre part ses différentes activités.

Les métas connaissances pour la représentation des connaissances du domaine d'application relèvent de la méta base chargée de contenir les structures des bases métiers. Comme ces bases sont utilisées pour l'exploration des contenus des bases et pour des analyses décisionnelles, les connaissances ou les informations contenues dans la méta base portent sur les attributs nécessaires pour ces explorations et analyses.

(a) Au niveau 1 on trouve les instances des deux bases métiers étudiants et enseignants obtenues par extraction de données de l'entrepôt, éventuellement complétée par des données externes :

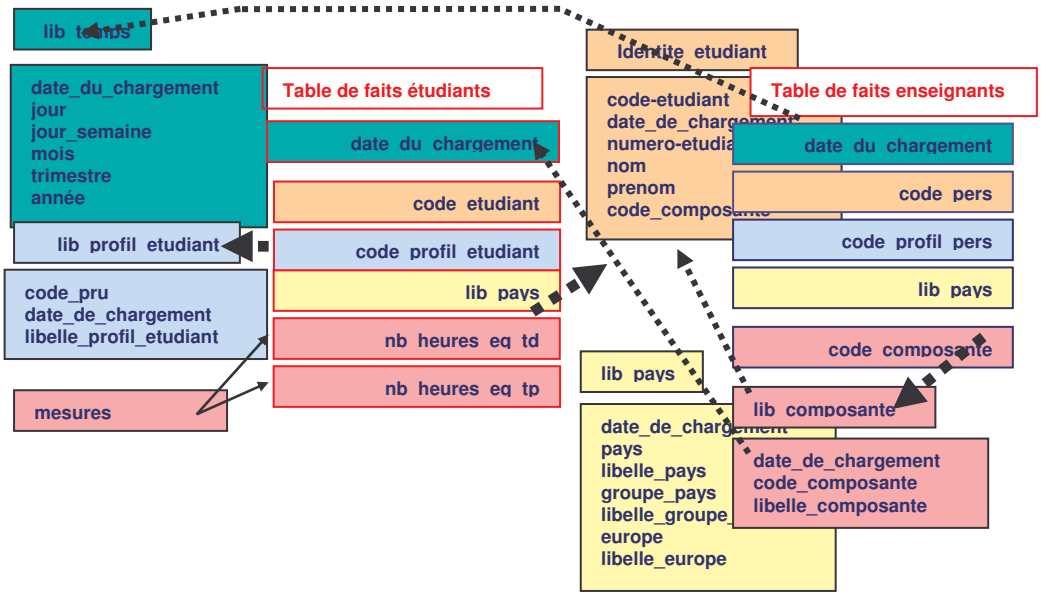

Figure 13. *Mise en relation de 2 tables de faits*

La constellation de faits (figure 13) permet de représenter plusieurs tables de faits partageant quelques tables de dimension Les tables de faits (étudiants) et (enseignants) peuvent être mises en relation par la dimension temps et géographie,

- Par la (figure 8), nous avons des bases métiers orientés « diplômes » pour les étudiants.
- La (figure 9) nous permet d'avoir des bases métiers orientées vers le chef de la scolarité et le directeur d'UFR pour prendre des décisions sur le déroulement des diplômes.

(b) Le niveau 2 permet de savoir comment extraire les données de la base de données initiale, vers la structure de l'entrepôt de données. A ce niveau :

- La (figure 5) représentant l'univers étudiant, permet de spécifier d'où vient l'information. Quels sont les liens entre la source d'information et la structure actuelle de l'entrepôt de données.
- La Classe d'objet individu (figure 6) illustre la structure du modèle utilisateur permettant de stocker les caractéristiques types de chacun des destinataires des bases : comment retrouver les informations de la base initiale vers la classe d'objet individu. Ce modèle est à adapter vers chaque destinataire : étudiant, enseignant, administratif.
- La (figure 7) est un schéma conceptuel : il donne pour un diplôme la conception d'une U.E. Pour un diplôme, il y a plusieurs unités d'enseignement. Ce niveau nous permet de savoir comment associer chaque unité d'enseignement au diplôme.

(c) Enfin au niveau 3, il faut définir la structure de l'entrepôt de données. Ce niveau permet d'élaborer la différence entre la structure pour l'étudiant, l'enseignant, l'administration :

- La méta structure de la base de données permettant de façon générique le stockage des schémas conceptuels des différentes bases métiers : étudiant, enseignant, administration (figure 2)
- De la même façon il faut définir la méta base permettant de stocker les structures des modèles utilisateurs intégrés aux différentes bases métiers.

La figure suivante donne un méta schéma permettant de stocker non seulement les structures de bases métiers comme tout méta schéma étendu ou UML mais aussi les métas modèles des acteurs. En effet cette méta structure est tout à fait générique.

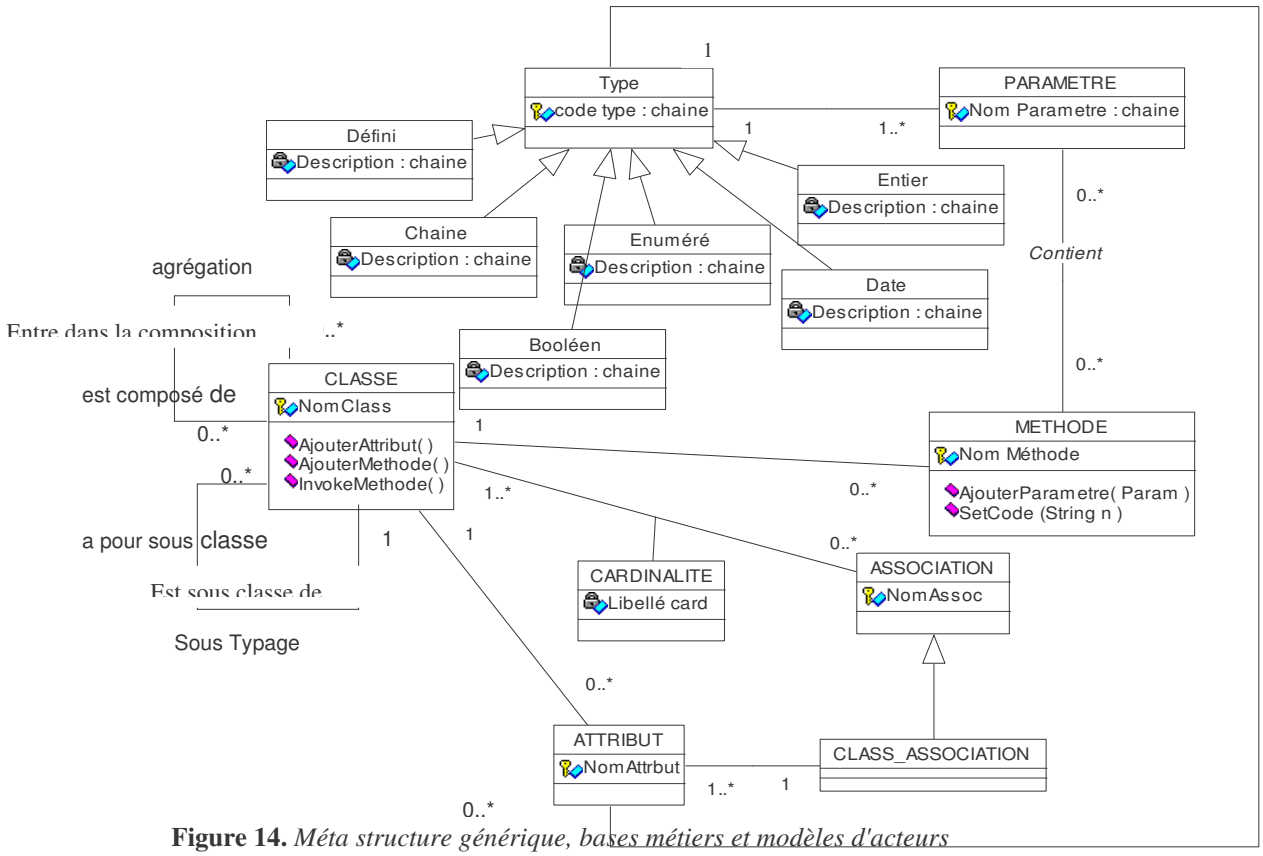

Figure 14. *Méta structure générique, bases métiers et modèles d'acteurs*

4. Conclusion

Les universités ont le souci de bâtir un système global d'information. Par ce papier, nous avons montré comment il est possible d'exploiter les bases de données existantes afin de concevoir un entrepôt de données qui intègre dès la conception les besoins des utilisateurs finals. Les bases métiers, orientées acteurs, permettent de suggérer des prises de décision. En analysant l'existant, nous avons mis en relief la structuration du système global d'information sous forme de brique. Cette construction facilite notre intégration pour apporter les éléments permettant de faire évoluer un système d'information en un système d'information décisionnel. La

modélisation fine des différents acteurs permettent d'élaborer les méta données des acteurs pour intégrer leur type, leurs besoins, leurs fonctions et leurs activités.
Nous élaborons les méta données propres à l'entrepôt de données en définissant les méta données structurelles et d'accessibilité propres au système de pilotage. Cela revient dans un S-IS à stocker parmi les métas-données du système, une représentation explicite de la structure des différentes bases métiers. Enfin, la métamodélisation nous permet de tracer l'évolutivité du schéma de l'entrepôt de données. L'aspect dynamique du schéma est dû au fait que le contenu de l'entrepôt est guidé par les besoins des utilisateurs. Il est donc nécessaire de prévoir l'évolutivité de l'entrepôt de données dès sa conception. Nos perspectives concernent la problématique du référentiel de méta données en tenant compte des certifications en vigueur.

5. Bibliographie

1. ADAE : Le répertoire des schémas XML de l'administration, http://www.adae.gouv.fr/article.php3?id_article=167.
2. Agence de modernisation des universités et des établissements, Séminaire Harpège, Présentation de l'univers Business Objects d'Harpège, 4 et 5 juin 2002.
3. Agence de mutualisation des universités, http://www.amue.fr/Amue/Default.asp.
4. Bouaka N. et David A., Modèle pour l'Explicitation d'un Problème Décisionnel : Un outil d'aide à la décision dans un contexte d'intelligence économique, IERA2003.
5. Boyer A., Nominé B., Managing new educative technology in a medium size university, In 20th ICDE World Conference on Open Learning and Distance Education. (Düsseldorf, Germany), 2001.
6. Cap Gémini Ernst & Young . Etudes préalables à l'élaboration d'un système d'information de gestion des établissements, 2003.
7. Chartron G., Les chercheurs et la documentation numérique : nouveaux services et usages, Paris, 2002.
8. David A., Modélisation de l'utilisateur et recherche coopérative d'information, 1999.
9. David A., Thiery O., Prise compte du profil de l'utilisateur dans un système d'information stratégique, in veille stratégique scientifique et technique - VSST'2001.
10. David A., Thiery O., Application of "equa2te" architecture in economic intelligence, 2002, http://ictei2002.loria.fr/papers/equate.htm.
11. David A., Thiery O., Prise en compte du profil de l'utilisateur dans un système d'information stratégique, VSST2001.
12. Desnos J.-F., Projet «Entrepôt de données», 2002, http://www.amue.fr/Telecharger/seminaire_pilotage_mars2002/J.F.Desnos.pdf.
13. Ducloy J., Cours IUT Paris 2002, http://dilib.inist.fr/~ducloy/iut.html.
14. 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/.
15. Franco J.-M., Le data warehouse : le data mining, Paris, 1997.
16. Goglin J.-F., La construction du datawarehouse : du datamart au dataweb, Paris, 2001.
17. Inmon-William.-H., Building the data warehouse, New York, 2002.
18. Kimball R., The data warehouse toolkit, John Wiley and Sons, 1996.

19. Kislin P., David A., Peguiron F., Caractérisation des éléments de solutions en recherche d'information : conception d'un modèle dynamique dans un contexte décisionnel, ISKO2003, Grenoble.

20. L'Essentiel d'Unified Modeling Language (UML), http://madchat.org/coding/other/CSI_UML_2003.pdf.

21. Manuel d'utilisation de l'infocentre PILOTAGE, http://ftp.amue.fr/documents_publics/apogee/II_doc_fonctionnelle/D_pilotage/MANINFO.doc.

22. Martre, H. Intelligence économique et stratégie des entreprises, Rapport du Commissariat Général au Plan, Paris, La Documentation Française, 1994.

23. Muller P.-A., Modélisation objet avec UML, Eyrolles 1997.

24. Muller P.-A., Représentation des vues d'architecture avec UML, http://magda.elibel.tm.fr/refs/UML/architecture.pdf.

25. Nataf J.-B., Structure de l'entrepôt de données de pilotage, 2001, http://www.cpu.fr/Telecharger/NatafP6StructureEntrepotDeDonnees.pdf.

26. Nominé B., ESUP portail : espace numérique de travail pour tous, Nancy 2004.

27. Peguiron F., Accès à l'information sur Internet, pratiques et tendances des utilisateurs : dans un contexte de documentation électronique, Université de Nancy-Metz, 2001.

28. Peguiron F., David A., Thiery O., Intelligence économique dans un cadre universitaire intégrant la modélisation de l'utilisateur, IERA 2003, Nancy, http://www.sciences.bu.u-nancy.fr/Parc/recherche/IERA2003.doc.

29. Peguiron F., Kislin P., Bouaka N., Activity-based classification of university actors for the construction of a domain-oriented data warehouse, SCI2003, http://www.sciences.bu.u-nancy.fr/Parc/recherche/SCI2003.doc.

30. Peguiron F., Thiery O. Modéliser l'acteur dans le système d'information stratégique d'une université, VSST 2004, http://www.sciences.bu.u-nancy.fr/Parc/recherche/VSST2004.doc.

31. Pottier, S., Mise en place de méthodes et d'outils pour le processus d'extraction de données en vue d'analyse décisionnelle : La méthode RADHE, Université de Nancy 2, 2002.

32. Revelli C., Intelligence stratégique sur Internet : comment développer des activités de veille et d'intelligence économique sur le web, 2000.

33. Rochfeld A., Morejon J., La Méthode Merise, Tome 3, Gamme opératoire, Editions d'Organisation, 1989.

34. Rumbaugh J., et al., Modélisation et conception orientées objet, Masson, 1995.

35. 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.

36. Teste O., Modélisation et manipulation d'entrepôts de données complexes et historisées, http://www.irit.fr/recherches/IRI/SIG/personnes/teste/these/intro.pdf.

37. Thiery O., David A., Modélisation de l'utilisateur : systèmes d'informations stratégiques et intelligence économique, Revue association pour le développement du logiciel (ADELI), 2002.

38. Thiéry O., Ducreau A., Bouaka N., David A., Piloter une organisation : de l'information stratégique à la modélisation de l'utilisateur ; application au domaine de la GRH, GREFIGE 2004.

39. Thiéry O, Support de cours recherches avancées en SIS DESS ACSI et SID de Nancy

40. Thiéry O, Support de cours SIS IUP Miage de Nancy