

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

**LORESA : Un système de recommandation d'objets d'apprentissage basé sur les
annotations sémantiques**

par
Sihem Benlizidia

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de Maître ès sciences (M.Sc.)
en informatique

Décembre, 2007

© Sihem Benlizidia, 2007.



Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé:

LORESA : Un système de recommandation d'objets d'apprentissage basé sur les annotations sémantiques

présenté par :

Sihem Benlizidia

a été évalué par un jury composé des personnes suivantes:

Jian-Yun Nie,	président-rapporteur
Esma Aïmeur,	directrice de recherche
Yann-Gaël Guéhéneuc,	membre du jury

Mémoire accepté le:

RÉSUMÉ

En cette dernière décennie, l'utilisation du e-Learning a connu une croissance considérable dans son déploiement dans différents domaines. Plusieurs institutions ont opté pour la centralisation de leurs ressources d'enseignement dans un seul emplacement : **les entrepôts d'objets d'apprentissage**. De cette façon, elles sont facilement accessibles aux enseignants ainsi qu'aux apprenants. Cependant, l'enrichissement continu du contenu de ces entrepôts pose le problème de la localisation de la ressource adéquate à la requête d'un utilisateur.

À cet effet nous proposons LORESA : Learning Objects Recommender System Based on Semantic Annotations. Ce système recommande des objets d'apprentissage à un utilisateur en se basant sur ses préférences et sur les annotations sémantiques des ressources. LORESA offre à l'auteur d'un objet d'apprentissage la possibilité de l'annoter sémantiquement, d'abord en choisissant dans une ontologie du domaine, trois concepts qui décrivent son objectif pédagogique, ensuite en donnant une valeur numérique à chaque concept pour en définir la *pertinence pédagogique* (une valeur sur une échelle de 1 à 5, où 1 désigne une faible pertinence et 5 une forte pertinence).

Ces nouvelles descriptions sont utilisées pour la génération des recommandations aux autres utilisateurs, qui sont alors invités à les évaluer. Ces évaluations aident à améliorer la précision des futures recommandations. Notre système utilise une technique hybride de recommandation : il utilise d'abord un filtrage basé sur le contenu en utilisant les préférences de l'utilisateur. Le filtrage collaboratif est ensuite utilisé en se basant sur l'évaluation des annotations sémantiques par le biais des votes des utilisateurs sur les objets recommandés.

Mots clés: objets d'apprentissage, Web sémantique, ontologies, systèmes de recommandation, metadonnées, annotations sémantiques, entrepôts d'objets d'apprentissage.

ABSTRACT

In the past decade, e-Learning initiatives have witnessed continued, sustained growth in their deployment in different domains. Different institutions have chosen to make learning resources readily accessible to educators and learners through learning object repositories ; but as more e-Learning repositories are created to store and manage learning object content, it becomes more difficult for the end users to find the right resource according to their needs.

We propose LORESA :Learning Objects REcommender system based on Semantic Annotation. This system helps recommending learning objects for a user based on his preferences and the resources semantic annotations. LORESA offers to authors of the objects, the possibility to semantically annotate them, first by choosing from domain ontology three concepts that define the learning object pedagogical objective, then by giving a value to each concept called the pedagogical relevance (a 5 value likert scale where 1 is for a poor relevance and 5 for excellent relevance).

These new objects descriptions are used for computing the recommendations for the other users, who are asked to evaluate them. These evaluations help improving the accuracy of the future recommendations. Our system uses a hybrid approach in the recommendation process :it first uses the content-based approach using the user preferences. Then it uses the collaborative approach based on the evaluation of the semantic annotations of the learning objects by the users ratings of the recommended objects.

Key Words :learning objects, Semantic Web, ontologies, recommender systems, metadata, semantic annotations, learning objects repositories.

TABLE DES MATIÈRES

RÉSUMÉ	iii
ABSTRACT	iv
TABLE DES MATIÈRES	v
LISTE DES TABLEAUX	ix
LISTE DES FIGURES	x
REMERCIEMENTS	xii
CHAPITRE 1 : INTRODUCTION	1
1.1 Problématique	1
1.2 Contribution	4
1.3 Plan du mémoire	5
CHAPITRE 2 : LES OBJETS D'APPRENTISSAGE	6
2.1 Généralités	6
2.2 Caractéristiques des objets d'apprentissage	7
2.3 Cycle de vie d'un objet d'apprentissage	8
2.4 Méta-données des objets d'apprentissage	10
2.4.1 Le schéma LOM de IEEE/LTSC	11
2.4.2 Le schéma IMS Meta-Data	12
2.4.3 Le schéma DCMI Meta-Data	12
2.5 Les entrepôts des objets d'apprentissages	13
2.5.1 Généralités	14
2.5.2 Quelques entrepôts d'objets d'apprentissage	15

2.6	Conclusion	18
-----	------------	----

CHAPITRE 3 : LES SYSTÈMES DE RECOMMANDATION DES OBJETS

	D'APPRENTISSAGE	19
3.1	Introduction	20
3.2	Les techniques de recommandation	20
3.2.1	Le filtrage collaboratif	21
3.2.2	Le filtrage basé sur le contenu	23
3.2.3	Le filtrage démographique	24
3.2.4	Le filtrage basé sur l'utilité	24
3.2.5	Le filtrage à base de connaissances	24
3.3	Les systèmes hybrides	24
3.3.1	Pondération (Weighted)	25
3.3.2	Commutation (Switching)	25
3.3.3	Technique mixte (Mixed)	25
3.3.4	Combinaison de caractéristiques (Features combination)	25
3.3.5	Cascade (Cascade)	25
3.4	Quelques systèmes de recommandation des objets d'apprentissage	26
3.4.1	Le système RACOFI	26
3.4.2	Le système COLDEX	28
3.5	Conclusion	30

CHAPITRE 4 : LES ONTOLOGIES : LEUR UTILISATION DANS L'AN-

	NOTATION SÉMANTIQUE	31
4.1	Introduction	31
4.2	Quelques définitions	32
4.2.1	Vocabulaire contrôlé	32
4.2.2	Taxinomie	32
4.2.3	Thésaurus	33

4.2.4	Réseaux sémantiques	33
4.2.5	Ontologie	33
4.3	Des ontologies pour la sémantique dans l'annotation	36
4.3.1	Introduction	36
4.3.2	L'annotation sémantique	37
4.3.3	Les standards	38
4.3.4	Quelques systèmes d'annotation sémantique	41
4.4	Conclusion	43
CHAPITRE 5 : CONCEPTION ET MÉTHODOLOGIE		45
5.1	Objectifs	45
5.2	Approche adoptée	47
5.2.1	Fonctionnement général	47
5.2.2	Fonctions principales accessibles aux utilisateurs	48
5.2.3	Techniques de recommandation utilisées	50
5.3	Méthodologie de LORESA	51
5.3.1	L'ontologie	54
5.3.2	L'entrepôt d'objets d'apprentissage	59
5.3.3	Le système d'annotation	59
5.3.4	Le système de recommandation	64
5.4	Architecture de LORESA	68
5.4.1	Composant de gestion de profil	69
5.4.2	Base des profils	69
5.4.3	Composant d'ajout d'objets d'apprentissage	72
5.4.4	Composant de recherche d'objets d'apprentissage	72
5.4.5	Composant d'annotation d'objets d'apprentissage	72
5.4.6	Composant d'enrichissement de l'ontologie	73
5.4.7	Composant de consultation de l'ontologie	73

5.4.8	Composant de recommandation des objets d'apprentissage	73
5.4.9	Composant d'évaluation des pertinences pédagogiques	74
5.4.10	Entrepôt des objets d'apprentissage	82
5.4.11	Ontologie du domaine	82
5.4.12	Table des annotations	83
5.4.13	Table des évaluations	83
5.5	Comparaison	83
5.6	Conclusion	87
CHAPITRE 6 : IMPLÉMENTATION ET VALIDATION		89
6.1	Implémentation	89
6.2	Utilisation de LORESA	90
6.2.1	LORESA : système de recommandation	92
6.2.2	LORESA : l'entrepôt d'objets d'apprentissage et le système d'an- notation	99
6.2.3	LORESA : l'ontologie du domaine	105
6.3	Validation	108
6.3.1	Contexte de la validation	108
6.3.2	Méthode	111
6.3.3	Validation des hypothèses	111
6.3.4	Conclusion	119
CHAPITRE 7 : CONCLUSION		120
BIBLIOGRAPHIE		124

LISTE DES TABLEAUX

5.1	Tableau comparatif de LORESA avec les outils d'annotation décrits en 5.4	63
5.2	Tableau des évaluations des utilisateurs de LORESA pour les quatre objets d'apprentissage	76
5.3	Tableau des Pertinences pédagogiques des quatre objets d'apprentissage .	77
5.4	Tableau comparatif du recommandeur de LORESA avec les systèmes décrits en 3.4	85
6.1	Tableau Descriptif du contexte de la validation de LORESA	109
6.2	Répartition selon la technique de recommandation utilisée et selon le profil	110
6.3	Résultats de l'étude de l'utilisabilité de LORESA	113
6.4	Résultats de la technique du filtrage par contenu	115
6.5	Résultats de la technique du filtrage collaboratif	116
6.6	Résultats de la technique hybride	116
6.7	Résultats des trois techniques : CT, CO et HY	117
6.8	Résultats de l'évaluation globale de LORESA	118

LISTE DES FIGURES

2.1	Cycle de vie de l'objet d'apprentissage [McGreal, 2004]	9
2.2	L'entrepôt MERLOT	16
2.3	L'entrepôt CAREO	17
2.4	L'entrepôt ARIADNE	18
3.1	Architecture du système Racofi [Boley et al., 2003]	27
4.1	Différents types d'ontologies.	35
5.1	Structure générale de LORESA	47
5.2	Structure de la classe <i>Domain</i> de l'ontologie de LORESA	55
5.3	Instances de la classe <i>Domain</i> de l'ontologie de LORESA	56
5.4	Structure de la hiérarchie des classes de l'ontologie de LORESA	57
5.5	Instance de la classe <i>SubDomain</i> de l'ontologie de LORESA	57
5.6	Instance de la classe <i>Category</i> de l'ontologie de LORESA	58
5.7	Instance de la classe <i>Concept</i> de l'ontologie de LORESA	58
5.8	Partie du fichier source de l'ontologie de LORESA en RDF/XML	59
5.9	Architecture de LORESA	70
6.1	Page d'accueil de LORESA	91
6.2	Page d'explication de la procédure d'enregistrement	92
6.3	Enregistrement de l'utilisateur <i>Jean</i>	92
6.4	Ressources recommandées à Jean lors de son authentification	93
6.5	Les annotations sémantiques des objets présentées à Jean	94
6.6	La recommandation de l'objet sélectionné	94
6.7	Consultation de l'objet	95
6.8	Explication de la recommandation	95
6.9	Choix de l'objet à évaluer	96

6.10	Choix des concepts à évaluer	97
6.11	Les nouvelles valeurs des pertinences pédagogiques	97
6.12	Profil de Jean	98
6.13	Nouveau profil de Jean	99
6.14	Modification du profil de Jean	99
6.15	Première étape d'enregistrement d'un objet d'apprentissage	100
6.16	Étape finale de l'enregistrement d'un objet d'apprentissage	100
6.17	Demande d'annotation de l'objet	101
6.18	Explication de la procédure d'annotation	101
6.19	Choix du champs sous-domaines de l'annotation	102
6.20	Choix de la catégorie de l'annotation	102
6.21	Choix des concepts et des pertinences pédagogiques	103
6.22	Recommandation possible de l'objet annoté	103
6.23	Recherche de ressources pédagogiques	104
6.24	Affichage des ressources pédagogiques répondants aux critères	104
6.25	Consultation des ressources du domaine <i>Computer Sytems</i>	105
6.26	Consultation de l'ontologie	106
6.27	Consultation des concepts	107
6.28	Ajout d'un concept à l'ontologie de LORESA	107
6.29	Répartition des participants de chaque catégorie selon le profil	110
6.30	Répartition des participants de chaque catégorie selon le profil	112
6.31	Comparaison des trois techniques	117
6.32	Évaluation globale de LORESA	118

REMERCIEMENTS

Je tiens à adresser mes sincères remerciements à ma directrice de recherche, professeure Esma Aïmeur, elle a su me guider lors de mon cheminement par ses conseils judicieux et ses questions pertinentes qui m'ont toujours aidés à améliorer les solutions proposées.

Par ce mémoire, je tiens également à remercier et témoigner de toute mon affection et mon attachement à ma famille : mon mari M'hamed, mes deux princesses Nour et Lina-Zahra, ma sœur Meriem, mes parents Mahmoud et Halima et mes frères Toufik et Imed. Je les remercie tous de leur soutien indéfectible et de leur présence à mes côtés.

Tous mes remerciements vont également à Soumaya, Zakia, Narimel et Leila pour leur amitié et leurs encouragements. Un grand merci à tous les membres du laboratoire Heron pour leur soutien.

Je remercie également les membres du jury qui ont bien voulu évaluer mon travail, les professeurs Jian-Yun Nie et Yann-Gaël Guéhéneuc.

CHAPITRE 1

INTRODUCTION

Les technologies de l'information ont eu sur le monde de l'éducation un impact indéniable, traduit par un profond changement dans les manières d'apprendre et d'enseigner. Cette mutation a commencé par l'utilisation des supports technologiques pour enrichir des cours traditionnels donnés en classe, donnant naissance aux premiers systèmes de *e-Learning*. L'internet a pour sa part offert la possibilité de délivrer avec une facilité considérable du contenu pédagogique sous forme électronique sans restrictions de lieu ou de temps. La flexibilité de cette forme d'enseignement a créé un engouement certain parmi différentes institutions pour l'adoption partielle ou totale de la formation à distance.

Des petites et moyennes entreprises, des établissements d'enseignement de différents paliers ont alors conçu des cours entiers pouvant être dispensés via le Web. Cette tendance a créé le problème de la gestion des ressources pédagogiques dont le nombre ne cesse de croître. Ces institutions veulent optimiser l'utilisation de ces ressources afin d'atteindre leurs objectifs de formation ou d'enseignement à un moindre coût. Pour ce faire, elles ont opté pour la centralisation de ces ressources dans un seul emplacement : **les entrepôts d'objets d'apprentissage**. Le défi majeur auquel doit faire face cette banque de ressources pédagogiques est de fournir la ressource la plus adéquate à la requête d'un utilisateur et réduire ainsi le temps et le coût de la création d'une nouvelle ressource. Ce défi devient de plus en plus difficile à relever car ces entrepôts sont constamment enrichis de nouveaux objets d'apprentissage. Dans la section suivante, nous allons définir la problématique concernant cet aspect et présenter notre approche

1.1 Problématique

Le Web a apporté des changements radicaux aux méthodes d'accès à l'information affectant ainsi tous les domaines de la vie courante. L'enseignement à distance est parmi

les domaines où cette métamorphose a été révolutionnaire. En effet nous assistons à un engouement qui se poursuit pour toutes les formes du e-Learning, que ce soit par le biais de cours dispensés à la demande via le Web tels que WebCT¹ ou de classes virtuelles où toutes les modalités de communication de l'enseignant et des apprenants sont virtuelles : courriel, messagerie instantanée, vidéoconférence, etc. Divers systèmes éducatifs tels que les systèmes tutoriels intelligents et les plateformes d'apprentissage en ligne sont de plus en plus utilisés pour répondre à des besoins de formation des institutions d'enseignement et de formation professionnelle. Bien que les approches concernant l'adoption d'une solution de e-Learning puissent être différentes, il n'en demeure pas moins que l'apprenant ainsi que les ressources pédagogiques à utiliser dans ce processus sont au cœur de toutes les approches [Konstantopoulos, 2001].

Les réseaux et les technologies fournissent un accès inconditionnel à des ressources pédagogiques et à du support d'apprentissage pour des apprenants n'importe où et n'importe quand. Cependant notons que l'engouement pour l'utilisation du Web dans un contexte d'apprentissage a conduit à une profusion de matériel pédagogique qui peut s'apparenter à une bibliothèque universelle. Mais contrairement à cette dernière où chaque document est clairement libellé et identifié de manière unique et non-ambigüe, le Web quant à lui ne présente comme facilité de recherche que celle assurée par les moteurs de recherche. En fait ces derniers, étant basés sur des recherches textuelles à base de mot-clé, sont peu fiables et encore moins efficaces. Ainsi, actuellement la problématique de la recherche d'information est entière dans le Web et l'est encore plus dans un contexte d'apprentissage.

Les informations disponibles sur le Web devenant de en plus nombreuses, la tâche de localiser l'information pertinente devient de plus en plus difficile. Parmi les solutions qui ont été proposées pour combler cette lacune, les systèmes de recommandations ont été utilisés pour aider à filtrer l'information et ne présenter que celle qui est pertinente à un utilisateur donné pour une tâche spécifique. Ces systèmes peuvent être basés sur l'analyse

¹<http://www.webct.com/webct>

du contenu de la requête ou sur un filtrage basé sur la collaboration de plusieurs usagers selon des critères déterminés.

Dans un contexte d'apprentissage, il est difficile pour un utilisateur (apprenant ou enseignant) de trouver la ressource qui réponde le mieux à son objectif. Un système de recommandations pourrait être d'un grand apport pour aider les utilisateurs à trouver les ressources pédagogiques les plus adéquates à leurs requêtes d'une part et à leurs profils d'une autre part. Soit le scénario suivant : Jean, un enseignant à l'université, se voit affecter en début d'année un cours sur les structures de données. Pour la préparation de la matière qu'il va dispenser, il se tourne naturellement vers le Web pour chercher les ressources qui traitent du sujet. Il est vite découragé par la masse d'informations que le moteur de recherche lui présente, il oriente alors sa recherche vers les entrepôts d'objets d'apprentissage.

La réponse à sa requête lui est rendue sous forme d'une liste de ressources qui satisfont aux critères qu'il a introduits. Il note que ces dernières sont présentées de façon inappropriée puisque les informations relatives au niveau de difficulté, au format technique et à l'apport pédagogique sont complètement occultées.

Envisageons maintenant que Jean dispose d'un système qui peut lui présenter pour chaque ressource trouvée : son degré de difficulté, son format technique et surtout *les concepts* qu'elle couvre. Ainsi s'il cherche une ressource traitant des arbres, il peut déjà à ce niveau savoir pour chaque ressource si elle couvre les concepts d'arbre AVL, d'arbre 2-3-4, d'arbre SBB ou arbre B. Ce système associe à chaque concept, une valeur numérique qui décrit *le degré de sa couverture* par la ressource. Cette valeur est initialement attribuée par l'auteur de la ressource, ensuite elle peut être évaluée continuellement par tous les utilisateurs de cet objet d'apprentissage. Les ressources trouvées sont présentées en ordre décroissant du degré de couverture et de façon *personnalisée* pour chaque utilisateur. En effet, la valeur du degré de couverture étant constamment évaluée par les utilisateurs, le système présente à Jean, pour chaque ressource, des valeurs évaluées par des utilisateurs qui lui *ressemblent*. Grâce à ces informations additionnelles qui permettent de mieux

décrire les ressources recherchées, Jean peut faire un choix éclairé en se basant sur les recommandations qu'il a reçues.

LORESA, est un système qui offre la possibilité de définir pour chaque objet son apport pédagogique sous la forme d'annotations sémantiques effectuées préalablement par l'auteur de l'objet. Ainsi pour n'importe quel utilisateur à l'instar de Jean, qui effectue sa recherche dans l'entrepôt, peut avant même de consulter l'objet connaître les *concepts* qu'il couvre et aussi *le degré de cette couverture*, évalué par des usagers qui lui ressemblent. Cette possibilité améliorerait considérablement l'exploitation des ressources pédagogiques sur le Web.

1.2 Contribution

Nous avons développé le système *LORESA*, un système de recommandation d'objets d'apprentissage, qui permet d'ajouter des ressources à un entrepôt, de décrire leurs apports pédagogiques sous forme d'annotations sémantiques, de rechercher des objets d'apprentissage et de recevoir des recommandations personnalisées. Les annotations sémantiques servent à décrire les concepts que couvre le contenu de la ressource, dans un vocabulaire contrôlé : l'ontologie du domaine que nous avons développée. L'utilisation d'une ontologie du domaine (sciences de l'informatique) garantit la cohérence de ces descriptions. Nous avons ensuite défini : *la pertinence pédagogique*, qui est une valeur numérique qui décrit la contribution pédagogique de l'objet à l'acquisition de chaque concept choisi dans l'annotation. Cette valeur n'est pas constante est peut être constamment évaluée par les consommateurs de la ressource. Nous avons par cette description additionnelle, comblé une lacune dans les schémas actuels de description des ressources pédagogiques qui est la dimension pédagogique. *LORESA* permet ensuite à un chercheur de ressource dans l'entrepôt, de recevoir des *recommandations personnalisées* basées sur ses *préférences* et sur celles des utilisateurs qui lui ressemblent. Ce dernier peut évaluer les recomman-

dations qu'il a reçues ou recevoir à la demande *l'explication* du processus de génération de ces recommandations. Nous avons par cette option, fourni à l'utilisateur un moyen supplémentaire pour l'aider dans son choix des ressources les plus adéquates à sa requête.

1.3 Plan du mémoire

Notre mémoire est structuré comme suit : dans le deuxième chapitre, nous définissons les objets d'apprentissages, leurs caractéristiques et leurs différents standards. Nous présentons également dans ce chapitre les entrepôts d'objets d'apprentissage en expliquant leurs principes et nous en présentons quelques-uns. Nous abordons ensuite le troisième chapitre par une introduction sommaire des fondements et lacunes des systèmes de recommandation et nous le concluons par la présentation de quelques systèmes de recommandation des ressources d'apprentissage. Le quatrième chapitre est dédié aux ontologies : leurs définitions ainsi que leurs fondements théoriques et nous nous intéressons particulièrement à leurs utilisations dans l'annotation sémantique, que nous traitons par la présentation de quelques systèmes. Au cinquième chapitre, nous présentons l'approche que nous avons adoptée pour la conception de LORESA et son architecture. Le sixième chapitre est consacré à la description de l'implémentation et de la validation du système. Finalement, nous concluons en présentant les perspectives futures de ce travail.

Dans la suite de ce mémoire, notons que nous allons utiliser, de façon interchangeable, le mot *ressource pédagogique* pour désigner **un objet d'apprentissage**, le mot *usager* pour désigner *un utilisateur* et le mot *métadonnée* pour désigner *une annotation* et le mot *vote* pour désigner *une évaluation*.

CHAPITRE 2

LES OBJETS D'APPRENTISSAGE

Dans ce chapitre, nous allons traiter essentiellement des objets d'apprentissage : leurs caractéristiques, leurs standards et leurs schémas de description. Nous allons également passer en revue les entrepôts de ressources pédagogiques en en décrivant quelques-uns.

2.1 Généralités

Bien que la dénomination d'objets d'apprentissage soit largement adoptée au sein de la communauté de la formation à distance, il n'existe pourtant pas encore de consensus quant à sa définition. Ainsi nous trouvons dans la littérature plusieurs descriptions pour ce concept : McGreal le définit comme une ressource numérique réutilisable qui peut être encapsulée dans un cours, un module ou un programme [McGreal, 2004]. Downes le définit comme : une entité numérique ou non [Downes, 2001] et Wiley comme une entité qui peut être indexée et utilisée dans différents contextes d'apprentissage [Wiley, 2001]. Paquette stipule qu'un objet d'apprentissage peut comprendre les matériels, les outils, les services, les personnes et les événements [Paquette, 2004] ; il peut être de taille et de nature différente, tels des textes, des documents audiovisuels, des didacticiels, des présentations ou simulations multimédias, etc. Ces ressources, qui ne sont pas nécessairement numérisées, incluent les imprimés et le matériel de laboratoire. Le matériel est utilisé par des apprenants mais aussi par d'autres acteurs. Un objet d'apprentissage est donc une entité digitale qui peut être utilisable et réutilisable dans différents contextes d'apprentissage. Prenons l'exemple d'un fichier vidéo destiné à expliquer le phénomène de l'apesanteur terrestre, qu'un enseignant utilise comme un support d'apprentissage dans un cours de géographie. Cette ressource peut être combinée avec d'autres ressources pour servir de démonstration dans un cours de physique afin d'expliquer le phénomène de l'accélération. Cependant une ressource numérique n'acquiert la dénomination d'objet d'apprentissage

que lorsqu'elle est utilisée dans un environnement d'apprentissage car elle peut avoir été créée initialement pour d'autres objectifs. Metros énonce dans [Metros, 2005] que pour qu'une ressource numérique puisse avoir la dénomination d'objet d'apprentissage, il est nécessaire qu'elle dispose des éléments suivants :

1. un objectif d'apprentissage ou l'énoncé de la compétence à développer,
2. une activité pratique et, finalement,
3. une stratégie d'évaluation.

2.2 Caractéristiques des objets d'apprentissage

Le coût de production des ressources numériques à vocation pédagogique sur le Web est considéré comme un important incitatif à la création des objets d'apprentissage. Prenons l'exemple d'un enseignant de mathématiques qui crée un cours sur la fonction $\sin(x)$ sur le Web, il l'agrément de graphiques, d'animations, d'exercices interactifs, etc. Cette création peut nécessiter l'intervention de graphistes et de concepteurs de pages Web, ce qui augmentera considérablement le coût de cette ressource. Si à chaque fois qu'un enseignant qui veut enseigner ce cours, doit lui créer une ressource pédagogique, le coût en sera multiplié. Considérons maintenant que cette ressource une fois qu'elle a été produite, est mise sur le Web à la disposition de divers acteurs de l'enseignement, cela réduira considérablement le coût et le temps inhérent à la préparation des cours. Ceci nous permet de déduire dans ce qui suit les caractéristiques de base d'un objet d'apprentissage.

– *Les objets d'apprentissage doivent être partageables*

Cette caractéristique implique que la ressource peut être distribuée via le Web. Même s'il existe dans la littérature des définitions d'objets d'apprentissage comme étant des ressources pouvant être numériques ou non, cependant quand nous parlons de ressources non-numériques telles que des livres, des cartes géographiques, il est impossible de les utiliser sous cette forme dans un cours en ligne.

– *Les objets d'apprentissage doivent être modulaires*

L'objet d'apprentissage n'est pas un cours entier mais une partie d'un cours. La création d'un cours en ligne nécessite donc l'assemblage de plusieurs objets en une seule entité. Cette caractéristique qu'on appelle aussi la granularité des objets d'apprentissage. Ceci rejoint la définition faite par Longmire : "*learning objects must be free standing, nonsequential, coherent and unitary*" [Longmire, 2000].

– *Les objets d'apprentissage doivent être interopérables*

L'interopérabilité est la caractéristique qui permet à des objets d'apprentissage produits sous différentes plateformes ou résidents dans différents entrepôts, d'être assemblés pour former un seul cours. Singh stipule que : "*the framework must allow content and their data to be exchanged and shared by separate tools and systems connected via the internet*" [Singh, 2000].

– *Les objets d'apprentissage doivent être localisés facilement*

Cette caractéristique est primordiale pour la réutilisation des objets d'apprentissage. Elle se traduit par l'utilisation d'une description appropriée du contenu de la ressource qui aide à sa localisation par des outils tels que les moteurs de recherche ou autres (voir section 2.3.3).

2.3 Cycle de vie d'un objet d'apprentissage

Un objet d'apprentissage en tant que ressource suit un cycle de vie qui comporte quatre phases. Durant ces phases l'objet passe de l'étape d'un ensemble d'objectifs conceptuels à l'étape d'une ressource utilisable via un support technologique. Ces phases sont expliquées en détails dans les sections suivantes (voir Figure 2.1).

La phase de conception Cette phase permet à l'équipe pédagogique d'identifier, les objectifs que les apprenants doivent atteindre, de recenser les contenus des enseignements, de penser leurs structurations et d'envisager le produit final au niveau du contenu.

La phase de production L'objectif de cette étape est de transformer un scénario pédagogique et des contenus en *objet pédagogique* aux modalités d'usage différentes selon

les modalités d'enseignement : en présentiel, en ligne, etc. C'est ici qu'il est nécessaire d'envisager le passage des contenus pédagogiques au multimédia, de prévoir les coûts de production et d'envisager de réutiliser les ressources, de projeter la possibilité d'actualisation des contenus à travers la séparation du fond et de la forme.

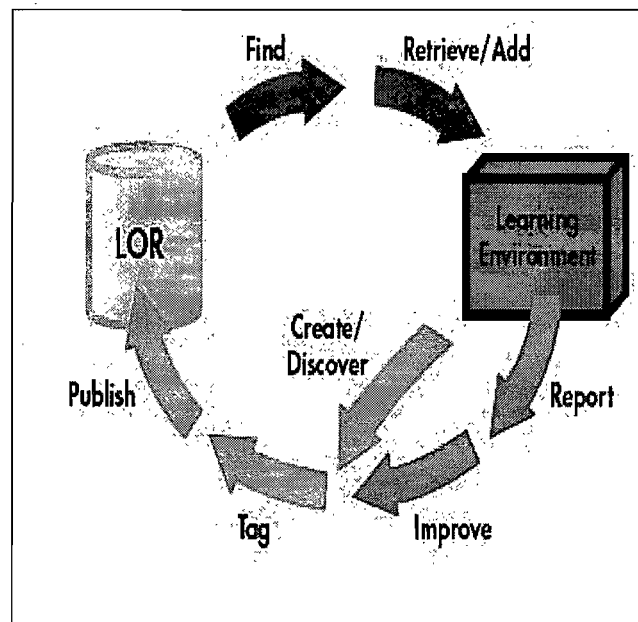


Figure 2.1 – Cycle de vie de l'objet d'apprentissage [McGreal, 2004]

La phase d'indexation L'indexation permet de créer une description pertinente pour chaque document pédagogique afin d'accomplir les objectifs suivants :

- La visibilité des objets d'apprentissage en les rendant accessibles.
- La réutilisation de ressources en intra pour créer de nouveaux objets (au sein d'un établissement ou d'un ensemble d'établissements).
- L'insertion dans des systèmes d'information ou de formation plus vastes, surtout si les descriptions sont aux standards internationaux.
- L'ouverture vers de nouveaux marchés : utilisation dans des contextes autres que l'enseignement.

La phase d'utilisation La phase d'utilisation permet de diffuser les objets pédagogiques via des dispositifs technologiques. Dans certains cas, il est possible d'exploiter, de traiter

et de collecter les résultats de l'activité de l'utilisateur. Il est nécessaire de rechercher l'interopérabilité du système pour être indépendant des plates-formes pédagogiques.

2.4 Méta-données des objets d'apprentissage

Une métadonnée est une donnée sur une donnée, plus précisément, c'est un ensemble structuré d'informations décrivant une ressource quelconque [Peccatte, 2002]. En formation en ligne, les métadonnées servent à décrire les objets d'apprentissage et comprennent des *descripteurs* qui permettent de les rendre plus facilement identifiables (accessibles par des mots-clés) et manipulables (interopérables, réutilisables, durables et adaptables). Pour cela l'utilisation de formats standards est indispensable et permet d'aboutir à une réelle réutilisation des ressources tel que le stipule Wiley : " *Without such standards, universities, corporations, and other organizations around the world would have no way of assuring the interoperability of their instructional technologies, specifically their learning objects*" [Wiley, 2000]. Afin de guider les divers praticiens du e-Learning que sont les concepteurs de contenu, développeurs de logiciels et utilisateurs dans une direction commune, différents groupes de travail ou comités ont été formés. Ces groupes permettent la collaboration entre personnes issues d'horizons différents afin d'établir des standards et des techniques accréditées et de recommander des pratiques dans le domaine du e-Learning. Les acteurs impliqués dans la standardisation sont nombreux. Il serait impossible de les présenter tous, notons cependant qu'actuellement, les trois standards les plus utilisés pour la description des ressources e-Learning sont : le schéma LOM de l'IEEE¹ (Learning Object Metadata), le schéma DCMI² et le schéma IMS Metadata³. Ces modèles de métadonnées décrivent comment le matériel d'apprentissage peut être décrit de façon interopérable. Tous les éléments des métadonnées nécessaires à la description d'une ressource peuvent être classés en plusieurs catégories, chacune offrant un point de vue dif-

¹<http://ltsc.ieee.org/wg12>

²<http://dublincore.org/>

³<http://www.imsproject.org/>

férent de la ressource. Par exemple, une image peut être définie par des métadonnées qui décrivent sa largeur, sa résolution, la date de sa création, etc. Les métadonnées d'un document texte peuvent comprendre : la taille du document, le nom de l'auteur, la date de sa création et un court résumé du contenu. Nous décrivons brièvement dans les sections suivantes, les éléments de base de chaque schéma.

2.4.1 Le schéma LOM de IEEE/LTSC

En juin 2002, le IEEE a approuvé la première version du standard Learning Object Metadata : LOM, qui est basé sur les spécifications issues des travaux des projets : ARIADNE⁴, SCORM⁵ d'ADL, AICC⁶. Il est graduellement devenu un standard de référence pour les systèmes d'éducation qui gèrent plusieurs objets d'apprentissage de différents types. Le standard de modèle de données LOM ou IEEE LTSC 1484.12.1 est seulement la première partie d'un standard multi-parties. Cette première partie contient un modèle abstrait (Abstract Model), des descripteurs qui sont utilisés pour décrire les objets d'apprentissage et qui ne sont pas concernés par la réalisation technique de ces éléments. Les éléments LOM seront gérés par différents formats incluant les tables SQL, les fichiers textes, les méta-balises HTML, etc. Cette réalisation technique du modèle abstrait dans un format spécifique, est appelée " Binding ". LOM décrit les ressources en utilisant un ensemble de plus de 70 attributs divisés en neuf catégories décrites par :

1. General
2. Lifecycle
3. Meta-Metadata
4. Technical
5. Educational
6. Rights

⁴[Http :www.ariadne.org](http://www.ariadne.org)

⁵[http ://www.adlnet.gov/](http://www.adlnet.gov/)

⁶[Http ://www.aicc.org/](http://www.aicc.org/)

7. Relation
8. Annotation
9. Classification

2.4.2 Le schéma IMS Meta-Data

IMS est un consortium formé par au moins 200 entités commerciales, gouvernementales et académiques. Il est issu du projet IMS-PROJECT qui a été initié en juin 1997 par le consortium EDUCOM consortium (actuellement EDUCAUSE). Son objectif majeur est d'établir des standards pour l'apprentissage en ligne parmi lesquels les standards des métadonnées des objets d'apprentissage. En 1998, IMS et ARIADNE ont soumis une proposition conjointe au comité LTSC de IEEE, cette proposition a formé l'élément de base du standard LOM. Ce dernier correspond au modèle IMS Learning Ressource Meta-data Information Model. Sa version actuelle est : IMS Meta-data Version 1.3, qui a été alignée au LOM 1484.12.1 2002.

2.4.3 Le schéma DCMI Meta-Data

Le Dublin Core Metadata Initiative Education Working Group est un groupe de travail du DCMI (Dublin Core Metadata Initiative). Son objectif consiste à discuter et de développer une proposition d'utilisation des méta-données Dublin Core pour la description des ressources éducatives. Le Dublin Core comprend 15 éléments dont la sémantique a été établie par un consensus international de professionnels provenant de diverses disciplines telles que la bibliothéconomie, l'informatique, le balisage de textes, la communauté muséologique et d'autres domaines connexes. Ses caractéristiques sont :

- Simplicité de création et de gestion.
- Sémantique communément comprise.
- Envergure internationale.
- Extensibilité.

Le Dublin Core est applicable à presque tous les formats de fichiers à condition que les métadonnées aient une forme interprétable à la fois par des moteurs de recherche et par des humains.

Notons que ces différents standards coexistent, et sont utilisés par divers acteurs du e-Learning. En fait, les méta-données de Dublin Core et de l'IEEE sont compatibles entre elles et interchangeables par simple mappage [Dublin Core Metadata Initiative 2004]. Ceci est aussi le cas pour les méta-données IMS et de l'IEEE. Un point commun qu'il est intéressant de relever entre ces différents standards est que chacun d'eux préconise l'utilisation d'un ensemble minimal de méta-données (telles que le titre, des mots-clé et le nom de l'auteur, par exemple), mais laissent à l'appréciation de l'utilisateur l'utilisation d'une multitude d'autres, permettant d'affiner la description au niveau souhaité.

En utilisant les standards décrits ci-dessus, la communauté du e-Learning est certes arrivée à un haut niveau de description du contenu qui permet son indépendance du contexte de sa création, facilitant ainsi son importation dans divers environnements. Cependant notons que tous les schémas utilisés ne disposent pas d'une composante sémantique *pédagogique*. Un niveau de description supérieur, rendant compte de l'aspect pédagogique, doit donc être ajouté aux objets d'apprentissage de manière à les rendre utilisables en tant qu'unités d'apprentissage. Nous proposons dans notre approche d'enrichir les standards actuels en définissant un sous-schéma de méta-données qui permet de décrire les concepts du domaine que couvre l'objet d'apprentissage en question et d'évaluer son degré de couverture pédagogique de ces concepts là.

2.5 Les entrepôts des objets d'apprentissages

L'utilisation de standards pour la description du contenu permet de développer des ressources réutilisables dans différents contextes d'apprentissage. Cependant le problème du dépôt et de l'accessibilité de ces objets, reste entièrement posé. Les concepteurs de cours veulent pouvoir déposer leurs objets sans que ces derniers soient inaccessibles car

ils seront stockés dans des architectures propriétaires. Le but de ce dépôt doit être de pouvoir les réutiliser, les mettre à jour, les archiver et les récupérer facilement. Il faut en outre qu'ils soient récupérables par des moteurs de recherche ainsi que par des humains. Ceci peut être réalisé en utilisant *les entrepôts d'objets d'apprentissage*. Nous présentons dans cette section quelques-uns de ces bibliothèques de ressources pédagogiques.

2.5.1 Généralités

L'entrepôt d'objets d'apprentissage peut être défini comme un serveur qui contient du contenu pédagogique organisé selon des métadonnées. Cette structure doit offrir la possibilité d'ajouter des objets d'apprentissage décrits par des métadonnées conformes à un schéma partagé mais aussi la possibilité aux autres utilisateurs de cet entrepôt de retrouver ces objets facilement. Les objets peuvent ne pas être physiquement tous sur le même serveur, cependant il est important que tout le contenu de l'entrepôt soit accessible aux utilisateurs de façon totalement transparente. Il existe plusieurs entrepôts d'objets d'apprentissage en ligne pour différents paliers d'enseignement, notons cependant que ceux qui ciblent le palier universitaire sont les plus nombreux. Comme nous l'avons vu en 2.1 et 2.2, un objet d'apprentissage peut être aussi simple qu'un paragraphe ou aussi complexe qu'un cours complet en ligne et être sous la forme de fichiers HTML et/ou texte, simulations, applets JAVA, Flash, etc. Ces entrepôts sont venus répondre à une demande de plus en plus accrue de matériel pédagogique réutilisable facilement localisable. Les moteurs de recherche en ligne basés sur la recherche textuelle ne donnent pas toujours des résultats pertinents à la demande d'une part et d'autre part, ils ne prennent pas en compte tous les formats d'objets : vidéo, Java applet, etc. Les entrepôts doivent ainsi fournir des interfaces utilisateurs adaptées à des recherches avancées, des soumissions d'objets, de commentaires et d'évaluations.

2.5.2 Quelques entrepôts d'objets d'apprentissage

Nous présentons dans cette section quelques entrepôts d'objets d'apprentissage parmi les plus utilisés.

2.5.2.1 MERLOT

MERLOT⁷ (Multimedia Educational Resource for Online Learning and Teaching) est un dépôt d'objets d'apprentissage créé en 1997, au centre d'apprentissage distribué de l'université d'état de Californie (CSU-CDL) dans le cadre du projet : "Authoring Tools and An Educational Object Economy (EOE) " avec la participation de partenaires universitaires, industriels et gouvernementaux. Son objectif majeur est la formation de communautés engagées à construire des bases de connaissances partageables de ressources d'apprentissage et à les évaluer pour en améliorer l'exploitation. Merlot offre les fonctionnalités de base d'un entrepôt : la recherche avancée, la soumission et l'évaluation de ressources d'apprentissage. Aujourd'hui, il compte un peu plus de 17000 ressources réparties en sept catégories : Arts, affaires, éducation, sciences humaines, mathématiques et statistiques, science et technologie et sciences sociales.

2.5.2.2 CAREO

L'entrepôt **CAREO**⁸ (Campus Alberta Repository of Educational Objects) a été réalisé conjointement par les universités de l'Alberta et de Calgary avec la coopération de BELLE (Broadband Enabled Lifelong Learning Environment) et de CANARIE⁹. Son objectif majeur est la création d'un fond de matériel pédagogique multidisciplinaire que les enseignants peuvent consulter via le Web. Le projet respecte les normes du schéma de métadonnées CanCore¹⁰ et s'inspire du projet MERLOT (voir 2.5.2.1). Comme l'entrepôt précédent, CAREO offre les possibilités de soumission, de recherche et d'évaluation

⁷<http://www.merlot.org/merlot/index.htm>

⁸

⁹<http://www.canarie.ca/about/index.html>

¹⁰<http://www.cancore.ca/fr/>

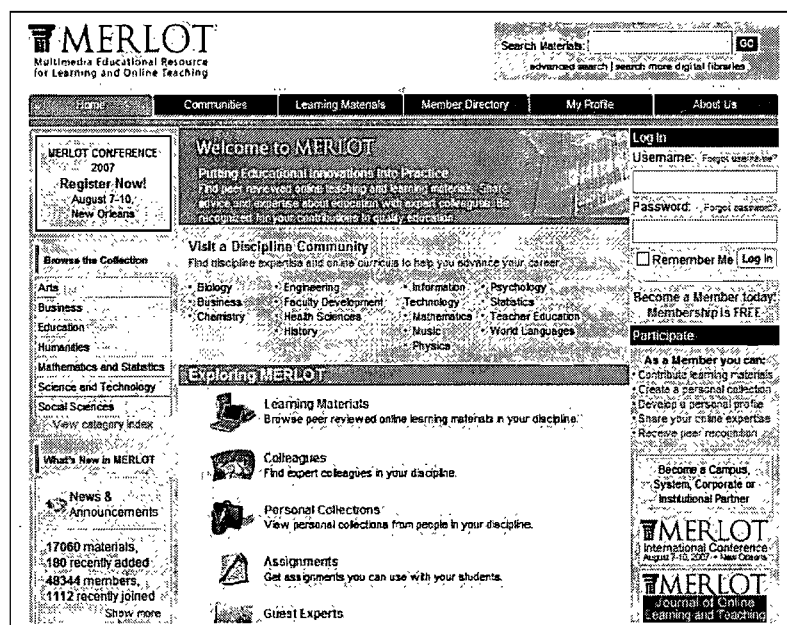


Figure 2.2 – L'entrepôt MERLOT

d'objets d'apprentissage. Son objectif majeur est la création d'un fond de matériel pédagogique multidisciplinaire que les enseignants peuvent consulter via le Web. Le projet respecte les normes du schéma de métadonnées CanCore¹¹ et s'inspire du projet MERLOT (voir 2.5.2.1). Comme l'entrepôt précédent, CAREO offre les possibilités de soumission, de recherche et d'évaluation d'objets d'apprentissage.

2.5.2.3 ARIADNE

ARIADNE¹² (Alliance of Remote Instructional Authoring and Distribution Networks for Europe) est une alliance qui regroupe 24 universités et grandes écoles de l'Europe et de la Suisse. Son principal objectif est de favoriser la production et les échanges de ressources pédagogiques multimédias entre les unités de formation de la communauté européenne (universités, écoles, entreprise). L'objectif majeur de cette alliance est d'optimiser l'utilisation des ressources pédagogiques créées par des organismes de formations sans

¹¹<http://www.cancore.ca/fr/>

¹²<http://www.ariadne-eu.org/>

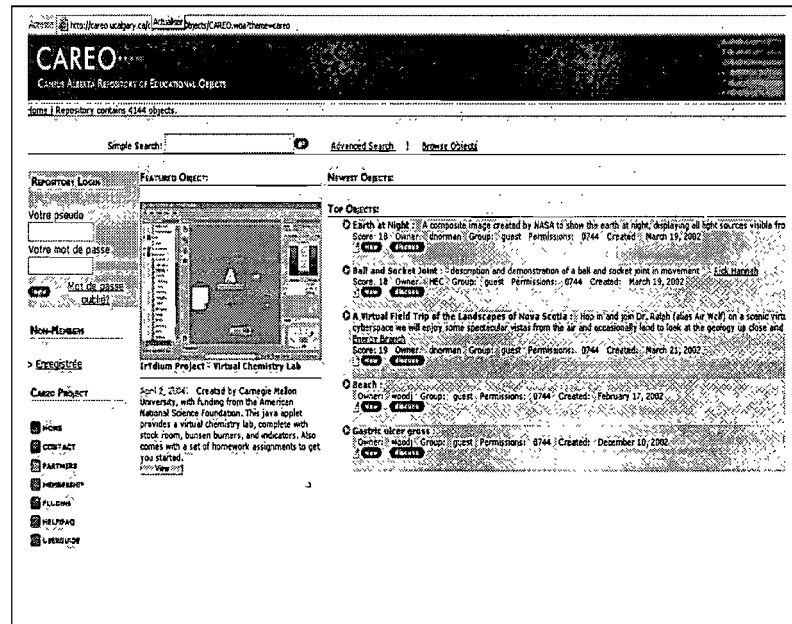


Figure 2.3 – L'entrepôt CAREO

référence commune à des normes en termes de mode d'apprentissage et d'ergonomie. Ces ressources sont très souvent distribuées via des serveurs Web où la mise à jour n'est pas contrôlable et où la recherche des contenus peut s'avérer trop aléatoire. Plusieurs outils ont été développés dans le cadre de ce projet et peuvent être utilisés conjointement à l'entrepôt, parmi lesquels :

- Knowledge Pool System : Gestion de la base distribuée et partagée des ressources.
- GenEval, OASIS, QuizCode, Sephyr, VideoClip generator : Outils de création de ressources pédagogiques, visant à faciliter la production par les enseignants et à assurer la qualité des documents.
- PHD generator : Indexation des ressources pédagogiques produites, avec des critères techniques mais aussi sémantiques et pédagogiques (méta-données).
- ALI (Ariadne Learner Interface) : Mise en uvre de ces curriculums auprès de populations d'étudiants, en mode présentiel ou à distance, via internet.

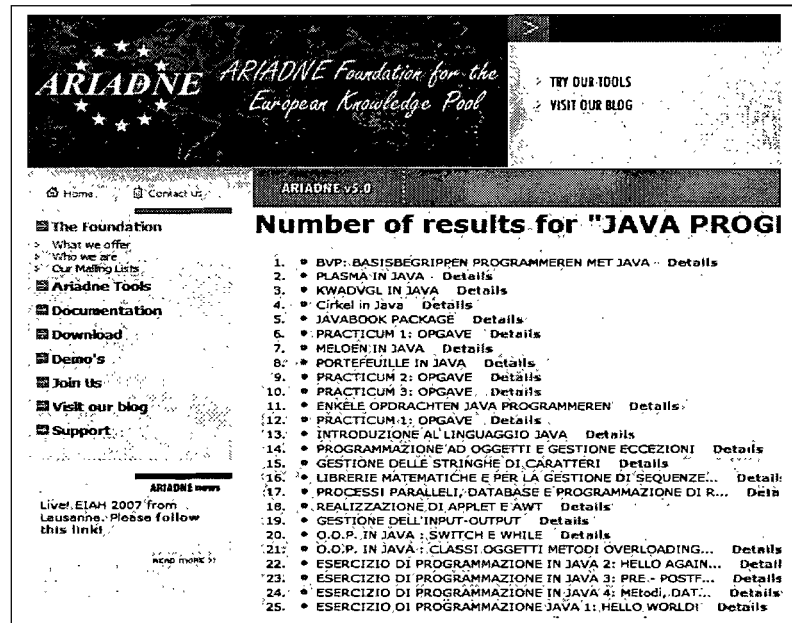


Figure 2.4 – L'entrepôt ARIADNE

2.6 Conclusion

Le développement (et la mise à jour) de matériel d'apprentissage a un coût. En outre, avec l'augmentation de la complexité de ces standards, le coût de leur application aux ressources d'apprentissage développées augmente lui aussi. Toutefois, l'ajout des métadonnées est une condition obligatoire au placement de ressources dans un entrepôt de ressources d'apprentissage, ce qui les rend réutilisables et, par là même, durables. Cependant que ce soit dans le contexte spécifique d'un entrepôt ou celui plus global du Web, ces ressources, dont le nombre ne cesse de croître, deviennent plus difficiles à localiser. L'aide à la recherche de matériel pédagogique d'intérêt devient alors bénéfique voire primordiale. L'utilisation des systèmes de recommandation à cet effet, n'est pas nouvelle et a fait l'objet de divers prototypes. Nous présentons dans le chapitre suivant quelques systèmes dédiés à cet effet, précédés par une présentation sommaire des systèmes de recommandations, leurs fondements et leurs lacunes.

CHAPITRE 3

LES SYSTÈMES DE RECOMMANDATION DES OBJETS D'APPRENTISSAGE

De plus en plus d'entreprises, d'institutions gouvernementales ou d'enseignement adoptent des solutions de e-Learning pour leurs besoins de formation, d'enseignement ou de consolidation. En effet la souplesse de ces solutions, la facilité de leurs mises en œuvre et surtout leurs coûts réduits porte à croire que cette tendance va se maintenir voire augmenter dans les prochaines années. Pour maintenir un coût acceptable de ces solutions, il est primordial d'assurer une exploitation optimale des ressources de base de ces systèmes en l'occurrence : les objets d'apprentissage en facilitant leurs réutilisations. Imaginons que le responsable de formation de l'entreprise "MEDIA-STAR " Jean, a à préparer un cours sur les systèmes SAP (Systems, Applications and Products for data processing) en un court laps de temps. Deux solutions s'offrent à lui :

- il peut concevoir son cours entièrement ('from scratch'), en utilisant du contenu qui traite des fondements théoriques avec des exemples pratiques et ce sous différents formats : texte, images, animations vidéo, simulations, etc. Cette opération nécessitera sûrement beaucoup de temps et d'efforts et donc un coût financier plus conséquent ;
- il peut recourir à des ressources pédagogiques déjà existantes soit au niveau de l'entreprise ou sur le Web. Cette solution lui permet de réutiliser des ressources existantes en les adaptant à son objectif de formation. pour répondre à un besoin spécifique. Cette solution semble à priori la plus rentable d'autant plus qu'il va utiliser des entrepôts d'objets d'apprentissage au lieu de faire sa recherche dans le Web entier.

À mesure que les solutions de e-Learning se démocratisent et gagnent en popularité, la deuxième solution semble de plus en plus prisée par les divers acteurs de la formation : entreprises, institutions d'enseignement, etc. Ceci a conduit à une prolifération de

ressources d'apprentissage qui à l'instar de n'importe quelle ressource sur le Web, sont de plus en plus difficiles à localiser, à évaluer dans un délai acceptable.

L'utilisation d'un système de recommandation dans ce contexte améliorerait considérablement la réutilisation de ces ressources, ainsi ils facilitent la tâche aux utilisateurs en les aidant à choisir la ressource qui convienne le mieux à leurs besoins et qui corresponde le mieux à leurs profils.

Dans ce chapitre, nous allons faire un bref état de l'art des techniques de recommandation les plus utilisées et nous présentons ensuite quelques systèmes de recommandation d'objets d'apprentissage.

3.1 Introduction

Un système de recommandation permet à un utilisateur donné d'avoir un service ou une ressource adaptés à son profil : ainsi il peut lui fournir des recommandations personnalisées ou le guider à choisir des produits (ou des services) qui lui sont utiles et qui répondent le mieux à son besoin [Burke, 2002]. Dans la section suivante, nous allons décrire les différentes techniques de recommandation.

3.2 Les techniques de recommandation

Selon Burke [Burke, 2002], il existe cinq principales techniques de recommandation : le filtrage par contenu, le filtrage collaboratif, le filtrage démographique, le filtrage basé sur l'utilité et le filtrage basé sur les connaissances, appelé aussi filtrage à base de connaissances. Il existe une classification plus sommaire réalisée par Adomivicius et Tuzhilin dans [Adomivicius et Tuzhilin, 2005] et qui classe ces techniques en trois grandes classes : le filtrage par contenu, le filtrage collaboratif et l'approche hybride qui utilise une combinaison des deux dernières. Ces techniques sont expliquées dans les sections suivantes.

3.2.1 Le filtrage collaboratif

Le filtrage collaboratif est la technique la plus populaire dans les systèmes recommandation. Son principe de fonctionnement est de calculer des similarités entre plusieurs utilisateurs en se basant sur les votes que ces derniers ont effectués sur des objets communs du système. Ces votes représentent le degré d'appréciation de l'objet par l'utilisateur [O'Sullivan et *al.*, 2002]. Le vote est donc une évaluation, qui peut-être textuelle où le système sollicite l'usager afin que ce dernier exprime son opinion par écrit. L'évaluation peut aussi être numérique, il s'agit d'attribuer une note au produit sur une échelle numérique (de 1 à 5, de 1 à 10, etc.). Le filtrage collaboratif comporte trois étapes principales, à savoir :

- L'utilisateur *vote* ou *évalue* différents objets proposés par le système.
- Le système calcule *la similarité* entre les différents utilisateurs en basant la fonction de similarité sur les votes. Si deux utilisateurs ont voté de façon semblable sur un produit donné, le système les considère comme proches, et prévoit des votes similaires de leur part sur d'autres produits.
- De nouvelles *recommandations* sont faites par le système, et ce en fonction des utilisateurs ayant le plus grand degré de similarité dans les votes.

Deux classes principales d'algorithme de filtrage collaboratif ont été identifiées par Breese et *al.* [Breese et *al.*, 1998] : les algorithmes basés sur la mémoire (*memory-based*) et ceux basés sur les modèles (*model-based*).

3.2.1.1 Algorithmes basés-mémoire (memory-based)

Les algorithmes basés sur la mémoire utilisent les profils utilisateurs afin de générer de nouvelles prédictions. Ces algorithmes s'appuient essentiellement sur deux techniques citées par Breese et *al.* [Breese et *al.*, 1998] : *la similarité vectorielle* et *la corrélation de Pearson*. Le modèle vectoriel se base sur une technique utilisée en recherche d'informations pour mesurer la similarité entre deux documents. Ainsi la similarité entre deux documents représentés par des vecteurs de mots-clés, est calculée par la formule du co-

sinus. Cette technique est utilisée dans le cas du filtrage collaboratif où l'utilisateur est assimilé à un document A . Les noms des objets sont considérés comme des mots-clés et les votes $v_{A,j}$ comme le poids donné par l'utilisateur A à l'objet j . Le poids est défini ici comme la fréquence de l'occurrence du mot-clé dans le document. Ainsi, la similarité vectorielle entre deux usagers A et B est calculée par la formule du cosinus présentée ci-dessous :

$$\text{cosinus}(A, B) = \sum_{j=1}^{nc} \frac{V_{Aj}}{\sqrt{\sum_j V_{Aj}^2}} \times \frac{V_{Bj}}{\sqrt{\sum_j V_{Bj}^2}} \quad (3.1)$$

nc : nombre d'items évalués en commun par A et B . v_{Aj} : vote de A pour l'item j . v_{Bj} : vote de B pour l'item j . La corrélation de Pearson, quant à elle permet de calculer la similarité entre deux usagers [Resnick et *al.*, 1994] par le biais de la formule qui suit :

$$w(A, B) = \frac{\sum_j (v_{Aj} - \bar{v}_A)(v_{Bj} - \bar{v}_B)}{\sqrt{\sum_j (v_{Aj} - \bar{v}_A)^2 \sum_j (v_{Bj} - \bar{v}_B)^2}} \quad (3.2)$$

Le système calcule la prédiction du vote sur l'objet j par l'utilisateur A en prenant compte sa corrélation avec tous les autres usagers ayant évalué des objets en commun. Elle est donnée par ce qui suit :

$$P_{(A,j)} = \bar{v}_A + \frac{\sum_{i=1}^n W(A, i)(v_{ij} - \bar{v}_i)}{\sum_{i=1}^n |w(A, i)|} \quad (3.3)$$

n : nombre d'utilisateurs présents dans le voisinage de A , ayant déjà voté l'objet j . v_{ij} : vote de l'utilisateur i pour l'objet j . \bar{v}_i : moyenne des votes de l'utilisateur i .

3.2.1.2 Algorithmes basés-modèles (model-based)

Cette technique est basée sur les profils des utilisateurs afin d'estimer ou d'apprendre un modèle, qui sera utilisé par la suite pour faire des prédictions pour un utilisateur donné. L'algorithme basé-modèle construit un modèle de l'utilisateur en se basant sur ses votes

antérieurs : il génère la probabilité qu'un utilisateur A effectue un vote v_{Aj} sur l'objet j en se basant sur des algorithmes d'apprentissage machine comme le modèle des clusters et le modèle de réseaux Bayésiens.

3.2.1.3 Les limites du filtrage collaboratif

Selon Burke dans [Burke, 2002], les deux principaux inconvénients du filtrage collaboratif découlent du fait que le calcul de la similarité ou de la corrélation entre les *utilisateurs* est basé sur les votes donnés à des *objets communs*. Ainsi pour un nouvel utilisateur n'ayant pas encore fourni de vote, il est difficile de fournir une recommandation. De même qu'il est peu probable de recommander un nouvel objet qu'aucun utilisateur n'a encore évalué. Ces deux problèmes sont connus sous le nom du *démarrage à froid* (*Cold Start*).

Un système utilisant le filtrage collaboratif a tendance à recommander les produits les plus populaires. Ainsi, si un utilisateur a des goûts différents de ceux de la masse des autres usagers, il devient difficile de lui prédire de bonnes recommandations. Il s'agit du problème de *Gray Sheep* [Claypool et al., 1999]. Le problème de performance, est aussi à considérer concernant les algorithmes basés utilisateurs. La complexité du temps de calcul augmente avec le nombre d'utilisateurs et le nombre d'objets [Sarwar et al., 2001].

3.2.2 Le filtrage basé sur le contenu

Dans un système qui utilise le filtrage basé sur le contenu, les recommandations sont générées en se basant sur les objets que l'utilisateur a appréciés auparavant. Le système lui propose ainsi des objets au contenu similaire. Ceci nécessite la représentation des objets par des caractéristiques pour pouvoir les comparer par la suite. Le filtrage basé sur le contenu est parmi les techniques les plus faciles à implémenter.

3.2.3 Le filtrage démographique

L'objectif des systèmes de recommandation basés sur le filtrage démographique est tout d'abord de créer des catégories d'utilisateurs en se basant sur leurs données démographiques. Les recommandations sont par la suite faites à un utilisateur en déterminant à quelle catégorie il appartient. Pour construire de telles catégories, il est demandé aux utilisateurs lorsqu'ils s'enregistrent de fournir certaines informations démographiques nécessaires qui peuvent s'avérer difficiles à recueillir.

3.2.4 Le filtrage basé sur l'utilité

Dans ce cas la recommandation faite à l'utilisateur est basée sur l'utilité calculée de l'item recommandé à l'utilisateur. Tout le problème réside dans la définition de cette fonction d'utilité. Elle est essentiellement basée sur des informations sur les items.

3.2.5 Le filtrage à base de connaissances

Cette technique de filtrage utilise des techniques de l'intelligence artificielle pour pouvoir inférer des recommandations à un utilisateur. Ainsi en connaissant les attentes d'un utilisateur par rapport à un item donné ainsi que ses préférences, le système peut grâce à un système à base de règles par exemple, déduire la façon dont un objet particulier peut satisfaire un utilisateur donné.

3.3 Les systèmes hybrides

Il est clair que toutes les techniques de recommandations présentées ci-haut présentent chacune des inconvénients. Selon Burke [Burke, 2002], afin d'améliorer les résultats des systèmes de recommandations, la combinaison de deux ou de plusieurs techniques dans un même système peut s'avérer intéressante, ces systèmes sont appelés : systèmes hybrides. Dans la section suivante, nous présentons de façon concise les différentes techniques hybrides.

3.3.1 Pondération (Weighted)

La recommandation finale est obtenue en combinant avec des poids prédéfinis les résultats ou les votes générés par différentes techniques de recommandation.

3.3.2 Commutation (Switching)

Cette technique permet de commuter entre différents modèles de recommandation et ce, selon la situation d'utilisation. Des critères de commutation sont ainsi définis par le système pour choisir à chaque fois de façon non ambiguë la technique de recommandation à utiliser.

3.3.3 Technique mixte (Mixed)

Les recommandations fournies à l'utilisateur sont calculées en utilisant plusieurs techniques [Smyth et Cotter, 1999]. Le meilleur exemple de cette technique est l'utilisation simultanée des filtrages collaboratif et par contenu pour pallier le problème du démarrage à froid inhérent au filtrage collaboratif, ainsi le filtrage basé sur le contenu permet d'obtenir des recommandations sur de nouveaux objets, et ce en se basant sur leurs descriptions respectives.

3.3.4 Combinaison de caractéristiques (Features combination)

Il s'agit de combiner les caractéristiques des informations qui sont fournies par les différentes méthodes de recommandation pour utiliser une technique unique sur l'ensemble des données. Cette méthode permet selon Burke [Burke, 2002] de produire des recommandations plus fiables

3.3.5 Cascade (Cascade)

Cette technique hybride est basée sur l'utilisation des résultats d'une première technique comme données en entrée pour une deuxième technique de recommandation. Ainsi,

des items sont choisis par la première technique qui sont ensuite affinés en utilisant la deuxième technique de recommandation.

L'avantage de cette méthode, réside dans le fait que la deuxième technique peut servir uniquement dans le cas de figure où les recommandations générées par la première ne sont pas satisfaisantes ou qu'elles nécessitent une discrimination additionnelle.

LORESA est basé sur cette technique, en effet il met en cascade une **technique de filtrage par contenu** et une **technique de recommandation collaborative** basée sur la similarité vectorielle pour prédire des recommandations d'objets d'apprentissage à un usager. Après avoir dressé les fondements théoriques des systèmes de recommandation, nous décrivons dans ce qui suit deux systèmes dédiés à la recommandation des objets d'apprentissage.

3.4 Quelques systèmes de recommandation des objets d'apprentissage

3.4.1 Le système RACOFI

Ce projet est le fruit de la collaboration du conseil national de la recherche (NRC) et du consortium Knowledge Pool Canada. Son objectif principal est de fournir des outils de recherche et de restitution d'objets d'apprentissage dans de grandes bases de données. À cet effet, la base de règles RuleML a été implémentée. Cette dernière est utile car elle est généraliste et interopérable et permet la construction des services donnés. RACOFI (Rule-Appling Collaborative Filtering), se base sur le traitement des métadonnées objectives et subjectives qui décrivent un objet. Il est construit autour de deux agents logiciels : une bibliothèque de routines de filtrage collaboratif appelée COFI¹ et une base de règles appelée OO jDREW². Les métadonnées objectives sont traitées par des règles écrites en XML, plus précisément en RuleML [Boley et *al.*, 2003]. La composante COFI génère des prédictions à partir desquelles OO jDREW génère des recommandations.

¹<http://savannah.nongnu.org/projects/cofi>

²<http://www.jdrew.org/ojdrew>

Pour sa validation, ce système a été utilisé dans un site canadien de recommandation de musique appelé RACOFI Music³. Un utilisateur enregistré peut parcourir des objets, les évaluer(voter) et en proposer de nouveaux. Les votes concernent cinq attributs sur une échelle de 0 à 10. Ces attributs sont : appréciation globale, lyriques, musique, originalité et production.

Les métadonnées *objectives* considérées dans cette architecture sont : le titre, l'auteur, la date de production et un lien hypertexte qui permet à l'utilisateur d'écouter le morceau. La base de règles applique des poids aux différents attributs qui sont configurables par l'utilisateur. Il existe également des règles spéciales qui servent par exemple à prévoir des votes pour tous les objets du même auteur si l'utilisateur actuel a apprécié au moins un objet du même auteur. Ce prototype peut être adaptable à un entrepôt d'objet d'apprentissage existant et supporter ainsi les évaluations et les recommandations multidimensionnelles.

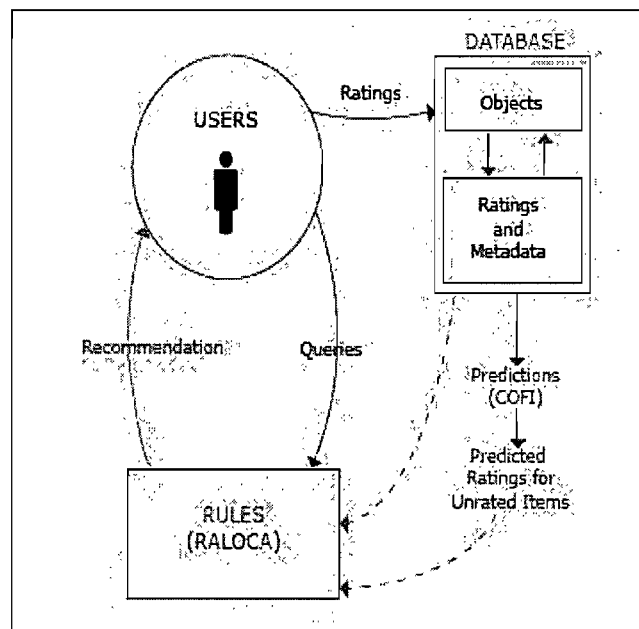


Figure 3.1 – Architecture du système Racofi [Boley et al., 2003]

Les concepteurs de RACOFI ont utilisé le filtrage collaboratif, introduit par Resnick

³ <http://racofi.elg.ca>

et *al.* [Resnick et *al.*, 1994]. Ils ont proposé une modification des algorithmes : Biais du moyen (BFM) [Herlocker et *al.*, 1999] et l'algorithme de Pearson, les deux algorithmes les plus cités et utilisés dans la littérature. L'algorithme proposé est "Scale and Translation Invariant (STI)". Cette modification leur a permis d'améliorer la précision d'une moyenne de 3%. Il est basé sur la factorisation de l'amplitude des moyennes des votes de tous les utilisateurs. Ils ont utilisé des implémentations en Java des algorithmes proposés dans leurs systèmes ainsi que XML pour la sauvegarde et la restitution des votes des utilisateurs ainsi que dans le calcul des prédictions.

3.4.2 Le système COLDEX

Le projet COLDEX [Baloian et *al.*, 2004] a été développé en 2004 dans le cadre d'une collaboration entre l'institut d'informatique de l'université du Chili et du département de génie électrique de l'université de Colombie. Son objectif principal est le développement d'environnements d'apprentissage distribués et collaboratifs pour des communautés d'apprenants en Europe et en Amérique du sud.

Il est constitué de l'interconnexion de réseaux de serveurs, où chaque serveur supporte une communauté d'apprentissage particulière. Des interfaces appropriées permettent un travail et un apprentissage collaboratifs. Les auteurs de ce projet ont développé un système de recommandation pour l'environnement décrit ci-haut. Les objets d'apprentissage étant de nature multimédia, il était impossible d'effectuer une analyse automatique du contenu, les auteurs ont donc opté pour la description du contenu de ces objets faite par des experts, dont le but est d'en déterminer la valeur didactique et de la représenter. En fait les auteurs proposent une méthodologie pour la description des objets d'apprentissage basée sur des techniques de filtrage collaboratif via un *vecteur de caractéristiques*. Ce vecteur contient donc les opinions des usagers ayant utilisé cet objet précédemment et évolue dynamiquement à mesure que plus d'utilisateurs accèdent à ce contenu.

La particularité de cette approche est que le vecteur de description d'un objet d'apprentissage est particulier à chaque utilisateur. Ainsi pour la construction du vecteur de des-

cription d'un objet pour un utilisateur particulier, les opinions des utilisateurs qui lui sont similaires auront un poids plus important. Cette méthode adaptative permet donc d'apprendre à partir de ces évaluations implicites et calibrer ainsi les préférences de cet utilisateur selon son comportement et ce pour présenter une prédiction plus précise à l'utilisateur.

Les auteurs ont utilisé quelques champs du standard LOM (voir 2.4.1) pour la description des objets d'apprentissage : format, langue, densité sémantique, taille, conditions d'installation et pré-requis. La description du document est donnée par : Doc_{jx} le j éme document de classe x . $FE_0(Doc_{jx})$ le vecteur d'évaluation initiale du document j , qui est une évaluation initiale de la contribution du contenu par rapport au contexte d'apprentissage, réalisée par un expert. Ce dernier donne une valeur entre 0 et 1, où 0 signifie que cet objet n'apporte aucune contribution au contexte d'apprentissage et 1 signifie qu'il apporte une contribution maximale. Pour chaque utilisateur k qui a utilisé et évalué le document au moins une fois, il existe un vecteur d'évaluation $FE_k(Doc_{jx})$, qui contient l'évaluation faite par l'utilisateur k de la contribution du document Doc_{jx} au contexte d'apprentissage. Le contenu du document Doc_{jx} est également décrit par le biais d'un vecteur qui reprend quelques éléments du standard LOM, notés par $CC(Doc_{jx})$ et Soit $CP_i(Doc_{jx})$ la contribution estimée de l'objet Doc_{jx} à l'utilisateur i générée automatiquement par le système selon la formule suivante :

$$CP_i(Doc_{jx}) = \frac{1}{N} * \sum_k FE_k(Doc_{jx}) * g(UP_i, UP_k) + FE_0(Doc_{jx}) \quad (3.4)$$

où N est le nombre des vecteurs $FE_k(Doc_{jx})$ en plus du vecteur $FE_0(Doc_{jx})$, la fonction g quant à elle détermine le degré de similitude entre l'utilisateur i et l'utilisateur k . Cette fonction tend vers 1 si les deux utilisateurs ont des profils rapprochés et tend vers 0 dans le cas contraire. Le système calcule donc le vecteur $CP_i(Doc_{jx})$ qui représente une estimation de l'évaluation du document par l'utilisateur. Après avoir consulté le document, l'utilisateur peut avaliser cette évaluation en déclarant que son évaluation est égale à sa

propre évaluation ou au contraire la rejeter auquel cas il faudrait introduire un vecteur avec de nouvelles valeurs qui sera pris en compte pour le calcul d'estimations futures pour d'autres membres de la communauté d'apprentissage.

La fonction g citée dans (3.4), évalue le poids de l'opinion d'un certain utilisateur à propos d'un document traduit par $FE_k(Doc_{jx})$ pour l'utilisateur pour lequel le vecteur $CP_i(Doc_{jx})$ est calculé. Cette fonction sert donc à accorder plus de poids à l'opinion d'un usager qui a un profil similaire à celui de l'utilisateur courant. Les auteurs utilisent pour le calcul de la corrélation statistique *la mesure du cosinus* et se sont inspirés des travaux [Savia et al., 1998], [Sheth, 1994] et [Yan et Garcia-Molina, 1994].

3.5 Conclusion

L'utilisation des systèmes de recommandation dans un contexte d'apprentissage répond de façon partielle à la problématique de réutilisation inhérente à l'exploitation des ressources d'apprentissage partagées. La majorité des systèmes de recommandation sont basés sur le partage d'informations entre différents utilisateurs, qui peuvent prendre forme d'évaluations, d'avis ou d'appréciations sur d'items. Toutefois, il est important de noter que les parties qui partagent ces informations ne s'accordent pas sur le sens accordé à ces informations, ce partage d'expériences peut vite devenir source de confusion et même induire en erreurs dans certains cas (un critère d'appréciation qui n'est pas compris de la même façon par deux utilisateurs, par exemple).

Pour que les systèmes de recommandations atteignent leurs pleins potentiels dans un contexte d'apprentissage, il devient nécessaire pour leurs utilisateurs d'utiliser *un vocabulaire* commun pour la description des critères d'évaluation de ces ressources. Leurs recherches en seraient ainsi facilités et leurs recommandations pertinentes à n'importe quel utilisateur. À cet effet nous introduisons dans le chapitre suivant les ontologies de façon générale ensuite nous décrivons leurs utilisations dans l'annotation de ressources.

CHAPITRE 4

LES ONTOLOGIES : LEUR UTILISATION DANS L'ANNOTATION SÉMANTIQUE

L'annotation sémantique des ressources pédagogiques, dans notre approche, nécessite l'utilisation d'un vocabulaire contrôlé et partagé par les acteurs du domaine. Ce vocabulaire se traduit par une ontologie qui doit être évolutive en fonction du besoin des utilisateurs. Dans notre travail, nous proposons de créer (ou enrichir) *une ontologie* afin de l'utiliser pour l'annotation (représentation des données) et la recherche de ressources pédagogiques (recherche d'informations). Dans ce chapitre, nous définissons les ontologies, leurs principes et fondements et nous nous attardons au rôle des ontologies pour l'annotation des ressources pédagogiques. Nous présentons alors quelques outils dédiés à cette tâche.

4.1 Introduction

L'organisation et la structuration des contenus est une tâche nécessaire à la création de n'importe quelle application. Ce besoin est devenu plus accru avec la rapidité de l'évolution de la masse d'informations dans tous les domaines. Les ontologies servent à la représentation des données échangées dans un domaine particulier afin de faciliter la communication interne au système informatique et externe entre les différents acteurs du domaine. Leur utilisation peut varier de la représentation des données à la recherche d'informations. Pour répondre à ces besoins, certains membres de la communauté scientifique (la communauté de l'IA, et dernièrement la communauté du Web sémantique) ont défini et utilisé le concept *d'ontologie*. Pour la communauté de l'IA (Intelligence Artificielle) : ce qui existe, c'est ce qui peut être représenté. L'ensemble des connaissances qui peuvent être représentées est alors appelé *l'univers du discours*.

4.2 Quelques définitions

Avant d'introduire la définition d'une ontologie, il est utile de décrire quelques concepts connexes.

4.2.1 Vocabulaire contrôlé

Un vocabulaire contrôlé est une liste de termes définis de manière consensuelle par une communauté afin de pouvoir décrire du contenu, et de rechercher l'information. Il est souvent utilisé pour les documents techniques ou plus généralement dans un domaine spécialisé. L'utilisation de ce vocabulaire contrôlé pour la description des ressources optimise leurs exploitations en facilitant leurs accessibilités aux utilisateurs potentiels. Cette facilité est obtenue par la délimitation du domaine dédié à la description et qui permet d'obtenir des résultats d'une recherche plus précis et pertinents par rapport à une description libre.

4.2.2 Taxinomie

C'est la structure la plus simple d'un vocabulaire contrôlé. En fait c'est une hiérarchie de termes, généralement organisés par le biais de la relation : spécialisation/généralisation. Ces termes peuvent être organisés selon d'autres types de relations telles que : la composition, l'imbrication, etc. Cependant un seul type de relation peut être représenté à la fois. [Texier, 2005].

Exemple de taxinomie :

- Mathématiques
 - Algèbre
 - Géométrie
 - Euclidienne
 - Sphérique

4.2.3 Thésaurus

Le thésaurus est une taxinomie augmentée de relations " horizontales " en plus de la relation verticale qui structure la taxinomie originale. Ces relations sont des liens du type " relatif à ", " voir aussi " ou " synonyme de ", etc et permettent d'élargir le champ d'étude ou de recherche. En ajoutant des relations du type " se traduit par... ", les thésaurus peuvent être utilisés pour construire des taxinomies multilingues. Néanmoins ceci n'est valable que si la modélisation choisie est universelle (indépendante de la langue et de la culture), une situation rare en dehors des domaines techniques et scientifiques.

4.2.4 Réseaux sémantiques

L'utilisation des graphes pour la représentation des connaissances en intelligence artificielle vient de l'idée de représenter graphiquement des concepts et leurs liens. Le premier outil proposé à cette fin est le réseau sémantique, introduit par Quillian [Quillian, 1968]. C'est un outil qui simule la représentation humaine de la mémoire. Il montre comment l'information pourrait être représentée en mémoire et comment on pourrait accéder à ces informations [Schwab, 2005].

Un réseau sémantique est donc une représentation de l'ensemble des connaissances qu'un individu se construit pour un domaine spécifique. Les éléments de connaissance ou les concepts sont représentés par des nœuds et sont reliés par des liens associatifs ou sémantiques [Habrant et *al.*, 1999]. Il est à noter que dans un réseau sémantique contrairement au thésaurus, un concept est défini de façon unique par des attributs propres aux nœuds et par des relations qui l'associent aux autres nœuds [Joubert et *al.*, 1992].

4.2.5 Ontologie

La définition la plus référencée et aussi la plus synthétique est sans doute celle de Gruber : "*une ontologie est une spécification explicite d'une conceptualisation*". Cette définition est étoffée dans [Gruber, 1992] où une ontologie est définie comme : "*un ensemble*

de définitions, de primitives, de représentation de connaissance spécifiques au contenu : classes, relations, fonctions et constantes d'objet". La même notion est également développée dans [Guarino, 1997a] où une ontologie est "*une théorie logique dont les modèles contraignent une certaine conceptualisation, sans la spécifier exactement*". Pour lui la définition de Gruber fait appel à la signification implicite d'une conceptualisation c'est pourquoi il la précise ; il considère les ontologies comme des bases de connaissances particulières. Pour [Grüninger et al., 1995] une ontologie "*est une description formelle d'entités et leurs propriétés, relations, contraintes, comportement*". Elle est simplifiée dans [Ikeda et al., 1999] où une ontologie est définie comme "*un ensemble de définitions de concepts et leurs relations à ne pas confondre avec un modèle qui est un ensemble d'instances de ces concepts.*" [Schoening, 1997] donne comme complément de définition "*une ontologie est un ensemble de spécifications de concepts compréhensible par une machine.*" Cette idée est renforcée dans [Devedzic, 1999] car une ontologie fournit la structure de base, l'armature autour de laquelle une base de connaissance peut être construite. [Mizoguchi, 1998] rajoute une définition qui provient du point de vue des systèmes à base de connaissance : *une ontologie est une théorie de concepts/vocabulaire, utilisée comme module des systèmes de traitement de l'information*. Les ontologies peuvent être classifiées en fonction de deux dimensions : leur niveau de détail et leur niveau de dépendance par rapport à une tâche particulière [Guarino, 1997b]. Nicola Guarino distingue alors plusieurs niveaux dans les ontologies (voir Figure 4.1) :

- *Les ontologies de haut niveau* contiennent les concepts généraux, communs à tous les domaines (temps, espace, objet, évènement). Les travaux de John Sowa et de Nicola Guarino tentent de formaliser une telle ontologie 'universelle' ;
- *Les ontologies liées à un domaine* particulier sont de deux sortes : soit elles contiennent le vocabulaire spécifique à un domaine bien défini et sont des spécialisations d'une ontologie de haut niveau, soit il s'agit d'ontologies de tâche qui contiennent l'ensemble des tâches réalisées dans un domaine donné ;
- *Les ontologies d'application* dépendent à la fois d'un domaine et d'une tâche.

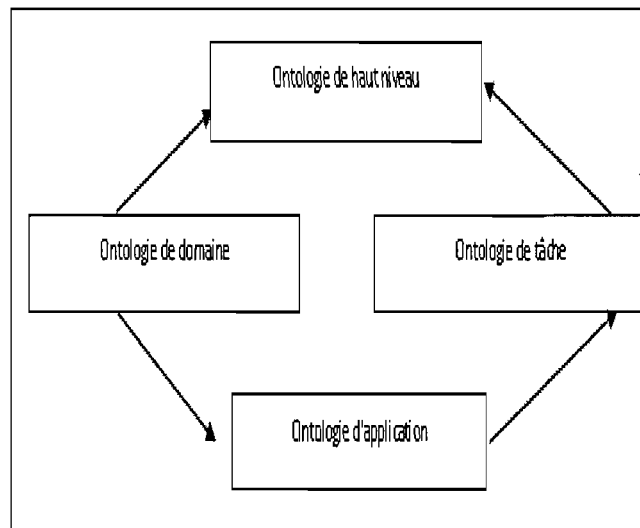


Figure 4.1 – Différents types d'ontologies.

Plusieurs ontologies ont été développées pour décrire des concepts généraux (ontologie de haut niveau) exemple : Cyc¹ qui est développée en utilisant le langage CycL. Une application qui utilise ses techniques est CycSecure. Cyc a la possibilité de construire des applications pour l'extraction des connaissances, pour la recherche intelligente, pour la traduction, etc. KR Ontology², qui est une ontologie de représentation des connaissances, a été dérivée de plusieurs sources telles que la logique, la linguistique, la philosophie et l'intelligence artificielle. Concernant des ontologies de domaines, on retrouve plusieurs exemples tels que Wordnet³ qui est une ontologie lexicographique pour la langue anglaise et UMLS⁴ qui est une ontologie médicale basée sur un réseau sémantique.

¹<http://www.cyc.com/cyc/technology/whatisyc>

²<http://www.jfsowa.com/ontology/>

³<http://wordnet.princeton.edu/>

⁴<http://www.nlm.nih.gov/research/umls/>

4.3 Des ontologies pour la sémantique dans l'annotation

4.3.1 Introduction

Le principe de l'utilisation des ontologies pour décrire des fragments de connaissance n'est pas nouveau [Gruber, 1993]. S'il s'est très vite répandu dans la communauté de l'Intelligence Artificielle, il semble aujourd'hui intéresser d'autres domaines tels que la vision par ordinateur ou la biologie. Pour souligner la diversité des domaines qui basent leur indexation sur les ontologies, nous pouvons citer les travaux importants de Karp en biologie [Karp, 1995], la bibliothèque électronique de l'université du Michigan qui utilise des ontologies pour représenter les méta-données qui décrivent les ouvrages et les articles qu'elle recèle [Weinstein et *al.*, 1997],[Weinstein, 1998] ou encore [Motta et *al.*, 2000] et [Domingue et *al.*, 1999] qui présentent un système de partage de documents journalistiques à l'intérieur d'une communauté.

Ces derniers travaux soulignent la contradiction qui peut exister à vouloir à la fois rechercher ce que nous appelons l'économie et la réutilisabilité. Les auteurs mettent en avant l'intérêt d'un support ontologique associé à un outil d'aide à l'indexation pour favoriser l'économie d'indexation. L'intérêt de se baser sur une ontologie partagée et réutilisable, pour réaliser une indexation avec le formalisme des graphes conceptuels est présenté dans [Martin et *al.*, 1996]. Enfin, même s'ils ne mentionnent pas explicitement le terme ontologie, Brusilovsky et *al.* soulignent l'importance d'indexer les unités pédagogiques avec des concepts du domaine [Brusilovsky et *al.*, 1996]. Chaque concept impliqué (appelé spectre) peut également représenter le rôle de l'unité qualifiée. **LORESA** est basé sur ce principe : nous proposons d'indexer les objets d'apprentissage par les concepts qu'ils couvrent en s'appuyant sur une ontologie du domaine. Les fondements de notre approche sont décrits plus en détails dans le chapitre 6. Nous donnons dans la suite du chapitre quelques définitions pour bien comprendre notre approche dans l'annotation des ressources et nous le concluons par la description de quelques outils dédiés à l'annotation sémantique des ressources sur le Web.

4.3.2 L'annotation sémantique

L'annotation sémantique revient à enrichir une ressource par des informations qui respectent *un vocabulaire particulier*. Ces informations sont appelées métadonnées et servent à décrire la ressource en question. Le vocabulaire que doivent respecter ces annotations est basé sur des ontologies. L'objectif de l'annotation sémantique d'une ressource est d'en optimiser l'exploitation automatique. Les métadonnées ajoutées servent à arriver à une meilleure catégorisation et à une indexation plus pertinente du contenu en levant toute ambiguïté sur son sens. Cette vision est défendue par les promoteurs du Web sémantique qui prônent une utilisation massive des ontologies pour la description du contenu du Web afin d'en optimiser l'exploitation. Ils proposent l'utilisation d'outils qui peuvent annoter le contenu des pages web, quels que soient leurs types, de manière manuelle ou automatique. La méthode automatique s'appuie sur des techniques de recherche textuelle et d'apprentissage machine pour l'ajout de ces métadonnées [Reeve et al., 2005]. Quant à la méthode manuelle elle se base sur l'intervention humaine pour le marquage sémantique du contenu à annoter. Cependant, deux problèmes présentent un réel défi à une large adoption de cette vision dans le Web. Le premier concerne l'utilisation de plusieurs ontologies qui couvrent le même domaine mais qui peuvent utiliser un vocabulaire différent pour un même concept. Cette problématique d'alignement fait actuellement l'objet d'un grand intérêt de la part de la communauté du Web sémantique. Le second problème est relié à l'interopérabilité des annotations qui utilisent des ontologies décrites par des langages différents, deux solutions peuvent être préconisées. La première concerne l'adoption massive d'un langage spécifique qui serait de ce fait un standard de facto. La seconde concerne le développement d'outils qui jouent un rôle de traducteurs entre les différents langages pour permettre une véritable interopérabilité. Nous pensons qu'à long terme, la première solution semble être la mieux adaptée. Nous introduisons dans la section suivante, quelques standards dans la description des métadonnées et des ontologies dans le Web.

4.3.3 Les standards

Pour la réalisation de la nouvelle vision du Web, il est évident que l'adoption massive de standards dans la présentation des documents et dans la formalisation de la connaissance est nécessaire voire indispensable. Les efforts du W3C⁵ se concentrent sur la standardisation des langages des ontologies et ceux de description des ressources. Nous allons présenter dans ce qui suit les principales recommandations du W3C dans ce domaine sont RDFS et OWL.

4.3.3.1 RDF et RDFS

RDF (Resource Description Framework) est un modèle, associé à une syntaxe, dont le but est de permettre à une communauté d'utilisateurs de partager les mêmes métadonnées pour des ressources partagées. Il a été conçu initialement par le W3C pour permettre de structurer l'information accessible sur le web et de l'indexer efficacement. La construction de base en RDF est le triplet sujet, objet, prédicat où :

- le sujet représente la ressource sur laquelle porte la déclaration, et c'est toujours un indicateur de ressource uniforme (ou adresse URI) ;
- le prédicat est le nom d'une propriété de la ressource (tel que le nom d'un champs d'enregistrement de base de données), et c'est toujours une adresse URI ;
- l'objet est la valeur de cette propriété, qui peut être une adresse URI ou un littéral (texte, nombre, date, etc.)

Le modèle décrit ci-dessus définit un moyen simple de décrire les relations entre des ressources en termes de propriétés (prédicat) et de valeurs (objet). Ces propriétés peuvent être considérées comme des attributs des ressources et correspondent alors au couple (attribut, valeur). Les propriétés RDF peuvent également être des relations entre des ressources. Ce modèle de donnée ressemble donc au modèle entité-association, cependant il n'offre aucun mécanisme pour la déclaration de ces propriétés ni pour la définition de ces rela-

⁵www.w3.org

tions. Ce rôle est dédié au schéma RDF ou RDFS (RDF Schema). Le mécanisme offert par le schéma RDF fournit un système de types de base à utiliser dans les modèles RDF. Ainsi il définit les ressources et les propriétés suivantes : **rdfs :Class** et **rdfs :subClassOf** qui sont utilisées pour la définition de types spécifiques aux applications qui les utilisent. Ces types sont spécifiés par le biais du modèle de donnée de base de RDF : ressources et propriétés. Le langage de spécification de ce schéma est un langage de représentation déclaratif qui est moins expressif mais beaucoup plus facile à implémenter que les langages de calcul de prédicats tels que : CycL ⁶ et KIF ⁷. Il est construit essentiellement sur les trois éléments suivants : les classes, les propriétés et les contraintes.

Les classes de base définies dans le vocabulaire de RDFS sont décrites ci dessous. Chaque modèle RDF basé sur l'espace de noms de RDFS les inclut.

- **rdfs :Resource**. Toute entité décrite par une expression RDF est appelée ressource et est considérée comme une instance de la classe **rdfs :Resource**.
- **rdf :Property**. C'est un sous-ensemble de **rdfs :resource** dont les éléments sont des propriétés.
- **rdfs :Class**. Elle définit le concept générique de *concept* ou *type* or *catégorie*. Elle est donc similaire à la notion de classe dans les langages orientés-objet. Lorsqu'une nouvelle classe est définie dans un schéma, la ressource qui la représente doit avoir une propriété **rdf :type** dont la valeur est **rdfs :Class**.

Les contraintes sont des descriptions des limitations d'utilisation des propriétés et des classes dans les données RDF. Par exemple un Schéma RDF doit décrire les limitations sur les types de valeur valides pour une certaine propriété. Généralement, le langage XML est utilisé pour exprimer des modèles RDF, on parle alors de RDF/XML.

⁶<http://www.cyc.com/cycdoc/ref/cycl-syntax.html>

⁷<http://www-ksl.stanford.edu/knowledge-sharing/kif/>

4.3.3.2 OWL

Le web sémantique est une vision d'un Web où le contenu possède une sémantique explicite, pouvant ainsi être traité par des agents logiciels. Les fondements de cette infrastructure commencent à prendre place notamment par l'utilisation de XML et de RDF pour la représentation des données. Cependant ces deux formalismes, à eux seuls ne suffisent pour une description consensuelle de la sémantique des documents Web. Une ontologie est donc indispensable pour décrire formellement la sémantique de la terminologie utilisée dans les documents Web. OWL est un langage qui se veut être la réponse à ces préoccupations. Le langage qui sert à définir ces ontologies doit donc dépasser les descriptions sémantiques de bases permises par RDFS. En fait OWL vient compléter l'infrastructure du web sémantique qui est construite comme suit :

- **XML** fournit une syntaxe pour écrire des documents structurés dont la sémantique n'est pas formellement exprimée ;
- **XMLS** sert à limiter la structure des documents XML mais étend également XML avec les types de données : *datatypes* ;
- **RDF** est un modèle de données pour définir des objets : "ressources" et les "relations" qui peuvent exister entre eux. Il fournit ainsi une sémantique de base pour ces modèles qui peuvent être exprimés en XML ;
- **RDFS**, est un vocabulaire qui sert à décrire les classes et propriétés pour des ressources RDF avec une sémantique pour la généralisation de ces hiérarchies de classes ;
- **OWL**, est venu enrichir le vocabulaire fourni par RDFS, pour décrire des relations entre les classes plus complexes et un typage plus riche des propriétés ;

L'utilisation de ces technologies dans les outils du Web garantit à moyen et à long terme une adhésion de plus en plus forte de la part des utilisateurs car elle permet une interopérabilité entre différentes données et facilite la collaboration.

4.3.4 Quelques systèmes d'annotation sémantique

Dans cette partie nous décrivons quelques systèmes d'annotation sémantique de ressources Web basés sur une méthode d'annotation manuelle. Nous nous intéressons à ces outils en particulier car nous avons adopté une approche similaire pour l'annotation des objets d'apprentissage. Nous insistons particulièrement sur la présentation des standards utilisés, des manières de stockage et d'exploitation des annotations créées. Les outils que nous avons choisis sont : Open Ontology Forge (2003), SMORE (2004), RDFPic (2002) et Annotea (2001).

4.3.4.1 Open Ontology Forge [Kawazoe *et al.*, 2003]

OOF (Open Ontology Forge) a été développé à l'institut national d'informatique de Tokyo au Japon. Son principe de fonctionnement est le suivant : l'utilisateur introduit l'adresse de la page Web qu'il veut visiter, il sélectionne ensuite le texte qu'il veut annoter et le fait glisser vers la classe de l'ontologie chargée qu'il juge la plus pertinente pour ce texte. Cette action ajoute automatiquement une nouvelle instance au concept désigné. OOF permet également l'enrichissement de l'ontologie en créant de nouvelles classes et en les instanciant par le biais de l'annotation.

L'originalité de OOF tient du fait qu'il offre la possibilité de définir les coréférences entre des instances déjà créées et des occurrences différentes dans le texte (des pronoms qui réfèrent à l'instance créée). Il cherche automatiquement les mêmes instances dans le reste du texte et les ajoute automatiquement. L'ancre de l'annotation est définie par un XPointer⁸ et les ontologies sont définies en RDFS. OOF intègre un éditeur d'ontologies, un navigateur et un support d'annotation. Il est très intéressant dans le cadre d'une utilisation personnelle d'annotation puisque les textes annotés peuvent être sauvegardés localement sous format XML.

⁸<http://www.w3.org/TR/WD-xptr>

4.3.4.2 Smore [Golbeck et al., 2002]

Smore a été développé au sein du laboratoire Mindswap à l'université du Maryland. Cet outil offre un environnement mixte de création de page Web et d'annotation sémantique. Il permet d'annoter des pages Web en se basant sur des ontologies existantes qu'on peut charger sous la forme de concepts, de propriétés et d'instances. Il offre également des possibilités de modification, d'extension d'ontologies via un éditeur d'ontologies SWOOP⁹ qui permet la création et l'exploration rapides d'ontologies OWL. L'annotation dans Smore est simple, en fait son principe est similaire à celui de l'outil précédent (OOF). L'utilisateur commence par sélectionner du texte de n'importe quelle page Web (fragments HTML) et par simple action de copier-coller au niveau du concept le plus pertinent au niveau de l'ontologie, l'annotation est créée. Smore offre également la possibilité de créer de nouvelles classes ou instances et de les sauvegarder localement. À l'instar de (OOF), les annotations créées ne peuvent être vues que par celui qui les a créées à moins que les utilisateurs qui désirent partager leurs annotations utilisent les mêmes ontologies.

4.3.4.3 RdfPic [Lafon et al., 2002]

RdfPic est un outil d'annotation d'images développé par des membres du W3C. L'annotation se fait en ajoutant des métadonnées RDF à une image via un éditeur de métadonnées. Le format de photo supporté par cette application est JPEG. Les métadonnées ajoutées sont classées en trois schémas :

- le schéma Dublin Core DC¹⁰ : c'est un schéma général qui sert à identifier les travaux originaux comme des livres ou des articles. Il contient les champs : créateur, éditeur, titre et date de publication. La version utilisée est en format RDF version 1.1 ;
- le schéma Technical : ce schéma sert à capturer les spécificités techniques de la photo et de l'appareil utilisé ;

⁹<http://www.mindswap.org/2004/SWOOP/>

¹⁰<http://dublincore.org/>

- le schéma du contenu (Content) : il contient les mots-clés utilisés dans le champ sujet du schéma DC, parmi lesquels portrait, baby ou landscape. Il sert à décrire le sujet de la photo en utilisant ce vocabulaire contrôlé et peut contenir autant de mots clés prédéfinis que nécessaire.

Cette annotation exprimée en RDF est incluse dans la photo elle-même et est sauvegardée en tant que bloc de commentaire. Ceci est selon nous un inconvénient car d'une part cela limite la taille de l'annotation mais surtout cela ne permet pas les annotations multiples pour une même photo.

4.3.4.4 Annotea [Kahan et al., 2001]

Annotea est une plateforme développée dans le cadre du projet LEAD (Live Early Adoption and Demonstration) du W3C. Cette plateforme permet l'implantation d'outils d'annotation de pages Web bâtis sur l'infrastructure RDF. Annotea permet d'ajouter des métadonnées décrites en schéma RDFS à des documents structurés XML ou HTML. Ces annotations sont elles mêmes décrites par des métadonnées (méta-métadonnées) externes au document et sont sauvegardées sur un ou plusieurs serveurs d'annotation. Elles contiennent la date, le nom de l'auteur, le type de l'annotation, l'URI du document annoté et l'URI du corps de l'annotation. Annotea utilise plusieurs spécifications définies par le W3C comme la combinaison de RDF avec XPointer, XLink¹¹ ou HTTP. L'approche d'Annotea pour l'annotation sémantique est de la présenter sous une forme semi-structurée par des champs prédéfinis. Cependant le corps de l'annotation est du texte libre qui ne respecte aucune ontologie.

4.4 Conclusion

L'annotation sémantique des ressources pédagogiques permet un meilleur partage et une utilisation plus pertinente des informations qu'elles contiennent. À cet effet nous

¹¹<http://www.w3.org/TR/xlink/>

proposons dans le chapitre suivant *une approche ontologique* qui sera intégrée à notre système de recommandation pour la formalisation du contenu pédagogique dans le but d'améliorer le partage des informations. Nous voulons ainsi utiliser les **ontologies** pour **annoter sémantiquement** les objets d'apprentissage par rapport à leurs objectifs pédagogiques. Nous comblerons ainsi une des faiblesses des métadonnées des objets d'apprentissage à savoir *la dimension pédagogique*.

Pour ce faire nous proposons dans un premier temps *l'annotation* du contenu de l'objet en question par des concepts issus d'une ontologie du domaine qui représentent son objectif pédagogique. Ensuite l'auteur évalue numériquement la contribution du contenu de sa ressource à l'acquisition des concepts choisis. Ces annotations sémantiques serviront à recommander des objets d'apprentissage mieux adaptés au *profil* de l'utilisateur qui les demande.

CHAPITRE 5

CONCEPTION ET MÉTHODOLOGIE

Dans ce chapitre nous allons décrire les objectifs de notre système et l'approche que nous avons adoptée pour les atteindre. Pour ce faire, nous allons d'abord décrire notre méthodologie pour la construction de notre système ensuite nous présentons l'architecture de LORESA¹ et les choix conceptuels que nous avons suivis.

5.1 Objectifs

De nos jours la formation est un processus qui nous accompagne tout au long de notre vie. L'utilisation du e-Learning et plus particulièrement dans sa forme basée-Web gagne en popularité dans les institutions et entreprises issues des milieux académiques, des affaires et des gouvernements. Cependant, à mesure que cette solution se démocratise, le défi auquel font face les responsables de formation de ces institutions est de plus en plus difficile à relever. Ils doivent *élaborer des cours personnalisés* qui répondent à des *objectifs pédagogiques précis* et ce au *temps opportun*.

La première étape de la construction d'un cours qui répond aux critères cités ci-dessus, est de chercher si une ou plusieurs parties de cette ressource n'ont pas déjà été traitées auparavant. En effet le coût de la production d'une telle ressource étant élevée, il est plus judicieux d'adopter lorsque cela est possible la solution de *la réutilisation*. Un autre défi attend notre concepteur de cours qui est de même difficulté sinon de difficulté supérieure, le nombre de ressources pédagogiques disponibles sur le Web a considérablement augmenté ces dernières années et le problème de leur *indexation* et de leur recherche devient aujourd'hui crucial.

Pour parer à cette première difficulté, les responsables de formation préfèrent diriger leurs recherche vers des banques spécialisées de ressources pédagogiques appelées : *entrepôts*

¹<http://www-etud.iro.umontreal.ca/benlisis/LORESA/src/index-fr1.php>

d'objets d'apprentissage. Ces derniers utilisent les normes et standards de métadonnées éducatives comme LOM et IMS Meta-data (voir sections : 2.4.1 et 2.4.2) pour l'indexation de leurs ressources. Cependant cette indexation du contenu de ces ressources a ses limites et ne couvre pas totalement tous les critères de recherche [Duitama *et al.*, 2005]. Nous proposons par le biais de **LORESA** (Learning Object REcommender based on Semantic Annotations), d'étendre ces standards en intégrant une description sémantique réalisée par l'*annotation sémantique* des ressources basée sur une ontologie du domaine couvert. Pour faire acquérir à l'objet qui est décrit sémantiquement la dimension pédagogique, nous avons proposé l'attribution de valeurs numériques à cette annotation qui traduit le degré de couverture de cette ressource des concepts qui l'annotent. Nous expliquons plus loin les choix que nous avons adoptés pour la concrétisation de cette vue. Une ressource est ainsi représentée par des *concepts* qui décrivent la sémantique du contenu couvert, mais aussi le degré de cette couverture évaluée par l'auteur de cet objet.

Cette description additionnelle permet d'offrir des outils de recherche plus sophistiqués qui permettent de *personnaliser* la restitution des ressources selon les concepts demandés et selon le profil du demandeur. Ainsi ces recommandations sont basées sur des évaluations d'utilisateurs qui ressemblent à cet usager. **LORESA** est donc le couplage d'un système d'annotation sémantique et d'un système de recommandation hybride appliqué à un entrepôt d'objets d'apprentissage. Les objectifs que nous voulons atteindre avec notre système sont :

1. Annoter les objets d'apprentissage sémantiquement par le biais *d'une ontologie du domaine*.
2. Permettre une évaluation *collaborative* du *lien pédagogique* de l'objet aux concepts annotants pour le décrire de manière à aider à sa recherche.
3. Permettre une recherche d'objets basée sur les critères du contenu.
4. Permettre une recommandation des objets basée sur la pertinence pédagogique des concepts recherchés : **Concept Pedagogical Relevance : CPR**

5.2 Approche adoptée

Cette section est dédiée à la présentation du fonctionnement global et des techniques de recommandation utilisées dans LORESA.

5.2.1 Fonctionnement général

LORESA est le *couplage* d'un système d'annotation sémantique et d'un système de recommandation appliqué à un *entrepôt d'objets d'apprentissage* (voir figure 5.1). Avant de pouvoir exploiter notre système, l'utilisateur doit d'abord s'enregistrer dans notre système en fournissant des informations telles que : pseudonyme, mot de passe, adresse courriel, son rôle : enseignant ou apprenant, la langue préférée des ressources recherchées, etc. Lorsque un utilisateur préalablement enregistré se connecte, le système lui présente des ressources d'apprentissage qui répondent aux critères qu'il a introduits lors de son enregistrement. Cependant l'utilisateur peut vouloir rechercher d'autres ressources en introduisant d'autres critères de recherche. LORESA restitue à l'utilisateur les objets d'apprentissage qui se trouvent dans l'entrepôt et qui répondent aux critères introduits.

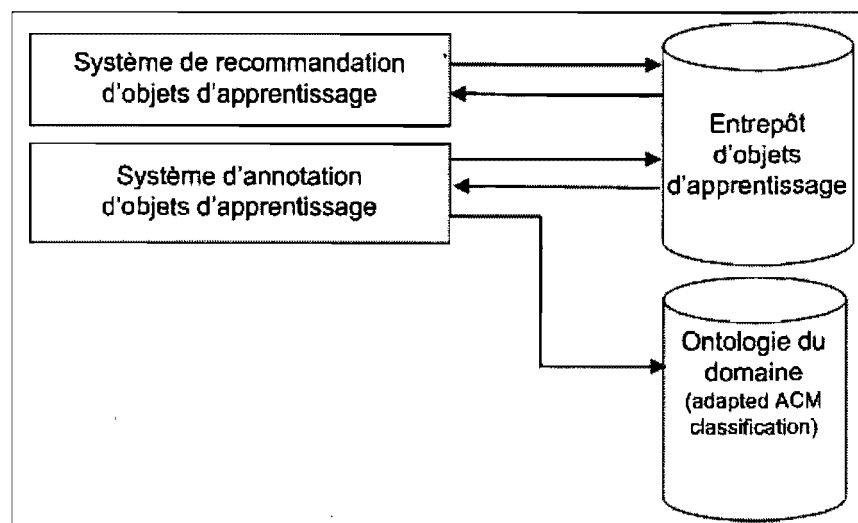


Figure 5.1 – Structure générale de LORESA

C'est à ce niveau qu'un nouveau choix s'offre à l'utilisateur, en effet le système lui propose de réaliser une *recherche sémantique*. C'est une option qui permet de réaliser une recherche plus approfondie qui se base sur les concepts couverts par les objets recherchés. Nous verrons cette recherche plus en détails dans les sections suivantes.

Si l'utilisateur est intéressé par certaines ressources il les choisit et peut recevoir à la demande les *recommandations* des usagers qui lui ressemblent sous formes de valeurs numériques. Quand l'utilisateur consulte les ressources qui lui ont été recommandées, il peut les évaluer en attribuant de nouvelles valeurs aux annotations sémantiques déjà évaluées par les utilisateurs. Cette nouvelle évaluation sert à donner des recommandations plus précises aux utilisateurs qui ressemblent le plus à notre utilisateur.

L'utilisateur peut également choisir d'enrichir l'entrepôt en soumettant un nouvel objet pour qu'il ajoute à l'entrepôt des objets d'apprentissage. A la fin de cette opération le système demande l'utilisateur s'il désire annoter l'objet introduit par le biais de l'ontologie chargée. L'annotation des ressources que nous proposons est une description pédagogique de cette ressource en définissant les concepts qu'elles couvrent plus précisément : un concept principal et deux concepts secondaires. Le processus d'annotation est expliqué dans les sections suivantes.

5.2.2 Fonctions principales accessibles aux utilisateurs

À la lumière de ce que nous avons introduit dans la section précédente, nous pouvons présenter les fonctions principales de notre système :

- *consulter des objets* d'apprentissage selon le domaine préféré : l'utilisateur introduit un ensemble de critères : domaine recherché, langue préférée, type de ressources, format, date de création, etc ;
- *ajouter des objets* d'apprentissage à l'entrepôt : cette option permet d'enrichir l'entrepôt d'objets d'apprentissage en soumettant un objet d'apprentissage décrit par des métadonnées ;
- *annoter sémantiquement des objets* d'apprentissage : à l'issue de l'opération d'ajout

d'un objet, le système demande à l'utilisateur s'il veut annoter son objet d'apprentissage. Cette option permet de définir le sujet traité par cette ressource en définissant les concepts qu'elle couvre. Cette description se traduit par l'attribution d'un concept principal et de deux concepts secondaires issus du domaine préalablement choisi pour cette ressource. c'est à ce niveau que l'ontologie prend toute son importance, en effet pour garantir la cohérence des annotations ajoutées, celles-ci doivent être prises parmi les concepts déjà définis dans l'ontologie. Si un concept n'existe pas, il peut être ajouté via l'option d'enrichissement de l'ontologie ;

- *rechercher des objets* d'apprentissage selon des critères définis : cette option permet à un usager de faire une recherche classique des objets d'apprentissage par le biais des métadonnées objectives ;
- *élaborer des recherches sémantiques* : ces dernières sont basées sur les annotations des objets et permettent d'aider l'utilisateur à choisir les ressources qui répondent le mieux à sa requête. Elles permettent de présenter à l'utilisateur les concepts couverts par l'objet en question et le *degré de couverture de chaque concept* représenté par une valeur numérique que nous avons définie par : CPR (Concept Pedagogical Relevance : pertinence pédagogique du concept annotant la ressource) ;
- *Obtenir des recommandations* sur les objets choisis : cette option permet de recommander à l'utilisateur courant les objets d'apprentissage en se basant sur son profil d'abord ensuite sur ceux des usagers les plus proches de lui ;
- *Obtenir des explications* sur les recommandations obtenues : Cette option est une des nombreuses contributions que nous avons réalisées par le biais de notre travail. Généralement le processus suivi par les systèmes de recommandation pour la génération des recommandations n'est pas expliquée aux usagers qui le perçoivent comme une boîte noire. LORESA ne se limite pas à une présentation des recommandations calculées à l'utilisateur, il permet de fournir à la demande de l'utilisateur une explication des recommandations présentées. Cette option est à notre sens, pertinente puisqu'elle permet à l'utilisateur de mieux évaluer les recommandations four-

nies ;

- *Evaluer les recommandations* produites par le système : l'utilisateur peut donner de nouvelles valeurs aux pertinences pédagogiques quand il n'est pas satisfait par celle générées par le système ;
- *consulter son profil* : objets ajoutés, objets annotés, objets évalués, etc ;
- *Modifier son mot de passe*, ses préférences, etc ;
- *Consulter l'ontologie utilisée* pour l'annotation des objets d'apprentissage ;
- *Enrichir l'ontologie utilisée* : cette option n'est disponible que pour l'expert dont le privilège est accordé par l'administrateur du système ;

5.2.3 Techniques de recommandation utilisées

La recommandation dans LORESA se base sur une technique hybride appelée la cascade qui combine le filtrage basé sur le contenu et le filtrage collaboratif. En effet, à chaque fois que l'utilisateur s'enregistre, le système lui recommande automatiquement des ressources qu'il a choisi en utilisant un filtrage à base de contenu. L'utilisateur peut approfondir sa recherche en demandant la recommandation des autres utilisateurs concernant ces ressources ou d'autres. Ceci est réalisé par l'application de la technique du filtrage collaboratif. Dans la première technique, la recommandation des objets d'apprentissage est basée sur les préférences de l'utilisateur, informations qu'il a fournies lors de son enregistrement et qui peuvent changer à sa demande. La deuxième technique implique la prise en compte des autres utilisateurs du système. Elle consiste à accumuler leurs votes sur les différents items (objets d'apprentissage), identifier ceux ayant voté les items en commun et offrir des recommandations en se basant sur des comparaisons faites entre l'utilisateur cible et ces utilisateurs.

Nous verrons dans la section 5.3.3 les détails de l'application de chacune de ces deux techniques.

5.3 Méthodologie de LORESA

La technologie actuelle de l'information fournit à l'utilisateur quel que soit son profil (apprenant ou enseignant) un accès facile et rapide à des ressources pédagogiques ainsi qu'à un support d'apprentissage. Cependant pour que cette facilité soit traduite par des gains en matière d'apprentissage ou d'enseignement, il est essentiel que *la recherche et l'utilisation* des ressources pédagogiques les plus adéquates soient aisées et performantes. En effet la multiplication des ressources pédagogiques sur le Web qui est certes un grand apport aux acteurs de l'enseignement, peut devenir un frein à leur bonne utilisation. C'est dans l'esprit d'offrir un moyen *d'optimiser* la recherche et l'utilisation des objets d'apprentissage que nous avons réalisé **LORESA**. C'est un système de recommandation des objets d'apprentissage qui est basé sur des annotations sémantiques de ces derniers pour aider au mieux l'utilisateur à trouver la ressource qui répond le mieux à sa requête. Pour atteindre les objectifs que nous avons assignés à LORESA et vérifier la validité de notre approche, nous avons suivi les étapes suivantes :

- Nous avons pris comme point de départ la classification de l'ACM des domaines de l'informatique [ACM, 1998] que nous avons traduite et enrichie en profondeur par l'ajout de concepts.
- La conception d'un entrepôt d'objets d'apprentissage qui fournit les fonctionnalités : d'ajout, de recherche et de consultation. Nous avons choisi de l'enrichir par des ressources d'apprentissage du domaine des sciences informatiques : algorithmiques, réseaux, intelligence artificielle, etc.
- La conception d'un système d'annotation sémantique des objets basée sur l'ontologie.
- La conception d'un système de recommandation des objets qui fournit selon le contexte soit des recommandations basées sur les préférences des utilisateurs qu'ils ont introduites lors de leurs enregistrements, soit des recommandations basées sur les évaluations collaboratives des ressources.

Nous avons voulu adresser deux faiblesses liées à l'utilisation des objets d'apprentissage pour la conception de cours, d'une part *la dimension pédagogique* de ces objets qui peut être soit mal exploitée ou non-définie. La deuxième faiblesse est liée *la personnalisation* de la présentation des objets recherchés. Nous voulons ainsi par par le biais d'une évaluation collaborative de couvertures pédagogiques des ressources pédagogiques en adapter les recommandations au profil du demandeur. Nous expliquons dans ce qui suit comment nous avons adressé ces deux lacunes.

La dimension pédagogique : parmi les objectifs visés par les efforts de standardisation des ressources pédagogiques, celui de la *réutilisation* est sans aucun doute le plus recherché. L'atteinte de cet objectif nécessite à notre avis une description précise du sujet traité par cette ressource. Même si les standards utilisés actuellement utilisent des champs spécifiques à cet effet, ils restent à notre sens insuffisants pour garantir une réutilisation optimale de ces ressources. Prenons l'exemple de Jean, un enseignant de réseaux qui cherche un objet d'apprentissage sur le protocole TCP. Il peut faire sa recherche dans un entrepôt en spécifiant dans sa requête quelques *mots clés* : modèle OSI, protocole TCP, etc. Les résultats de cette recherche vont être sans aucun doute importants et peu précis. Si cet enseignant avait la possibilité de spécifier de manière précise *les concepts* que doit couvrir la ressource recherchée, il est évident que le résultat serait plus pertinent. Nous proposons de combler ce manque en introduisant de nouveaux champs de description des objets d'apprentissage qui vont utiliser une description formelle du domaine choisi en l'occurrence : une ontologie. L'utilisation de l'ontologie garantit la consistance des descriptions et permet donc une utilisation optimale de cette ressource.

Cette description formelle est à notre sens insuffisante pour atteindre une bonne description pédagogique des objets d'apprentissage. À cet effet nous proposons *d'évaluer* le degré de couverture de la ressource des concepts annotants par des valeurs numériques. Soit la ressource *RESOI*, qui traite des réseaux de communication, plus précisément du protocole TCP. Le contenu de *RESOI* introduit plusieurs concepts relatifs au protocole cités tels : *Multiplexage/Démultiplexage, Acquiescement, Congestion*, etc. Toutefois, notons que

l'objet d'apprentissage en question, explique en détails le premier concept, il introduit de façon sommaire les deux autres concepts. L'auteur de la ressource évalue alors la couverture pédagogique assurée par *RESOI*, pour acquérir les concepts cités plus haut. Il décide alors que le concept principal couvert par son objet est : *Multiplexage/Démultiplexage* et lui accorde la valeur de cinq, pour son degré de couverture pédagogique. De la même manière il décide de prendre les concepts : *Aquittement* et *Congestion* comme concepts secondaires et leur accorde respectivement les valeurs de : 3 et 1 pour leurs degrés de couvertures. Ces informations renseignent les utilisateurs potentiels sur *l'apport pédagogique de RESOI* à l'acquisition des concepts sus-cités.

La dimension de la personnalisation : l'évaluation de l'apport pédagogique d'une ressource d'apprentissage réalisée par son auteur est subjective dans la mesure où ce qu'il peut juger lui comme un apport élevé, peut être jugé comme un apport moyen par un autre utilisateur. À cet effet nous proposons de permettre des *évaluations collaboratives* de ces degrés de couverture. Ceci permet d'attribuer à un même objet plusieurs valeurs de couverture qui dépendent du profil de l'utilisateur qui le recherche. Ainsi si Lina une enseignante de bases de données recherche des objets d'apprentissage qui traitent de certains concepts de la fibre optique, même si elle obtiendrait les mêmes objets que Jean, les degrés de couverture pédagogiques seraient sans doute différents car ils ont des profils différents. A chaque usager, LORESA présente des objets d'apprentissage avec des degrés de couverture pédagogique personnalisés affichés sous forme de *recommandations* faites par des utilisateurs qui lui ressemblent.

Pour pouvoir atteindre les objectifs que nous avons assignés à notre système, nous avons suivi une méthodologie incrémentale, ainsi nous avons commencé par construire notre ontologie qui est un élément d'une importance majeure dans notre approche. Nous avons ensuite enchaîné par la conception du système d'annotation et enfin le système de recommandation qui va utiliser les *outputs* du sous-système précédent (le système d'annotation) : les objets annotés et évalués pour calculer les résultats finaux à présenter à l'utilisateur : **des recommandations personnalisées**. Nous allons dans ce qui suit décrire et

expliquer nos choix conceptuels concernant ces trois éléments de notre architecture et les comparer à des systèmes similaires de la littérature pour faire ressortir notre contribution.

5.3.1 L'ontologie

La construction d'une ontologie consiste à la recherche de *concepts* généraux réutilisables, partageables et durables pour construire un modèle de connaissances capable d'aider des personnes à résoudre des problèmes [Mizoguchi, 2004]. L'utilisation des ontologies dans différents secteurs (biomédical, automobile, etc.) a connu une expansion due essentiellement à un souhait de s'orienter vers une vision d'un Web plus sémantique (voir chapitre 4). Le but commun, dans ces différents secteurs, est l'optimisation de la représentation des connaissances et leurs exploitation. Les objets d'apprentissage en tant que ressources partageables n'échappent pas à cette règle et doivent donc être décrits de manière à être *facilement retrouvés et utilisés*. C'est ce que nous avons voulu atteindre par le biais de la création de l'ontologie de concepts de LORESA. Ainsi en prenant comme point de départ la classification de l'ACM du cursus informatique, nous l'avons enrichie à un niveau plus élevé de granularité, et ce en étendant la description au niveau de concepts du domaine. Nous proposons donc une approche d'annotation des objets d'apprentissage qui a comme résultat trois concepts pondérés issus de l'ontologie du domaine. Parce que cette annotation sert à décrire formellement le sujet traité par l'objet d'apprentissage par le biais de concepts, le problème de l'enrichissement de l'ontologie doit être pris en compte. En effet, le volume de l'entrepôt d'apprentissage évolue dans le temps par l'apparition de nouveaux objets d'apprentissage qui nécessite la définition de nouveaux concepts pour les décrire. Dans notre système, nous proposons une solution pour enrichir l'ontologie du domaine. Notons que cet enrichissement ne comprend pas la suppression et la transformation de concepts. L'ontologie enrichie ici est un ensemble de termes reliés par la relation de spécialisation-généralisation.

Pour la construction de notre ontologie, nous avons utilisé **protégé**². C'est une plate-

²<http://protege.stanford.edu/>

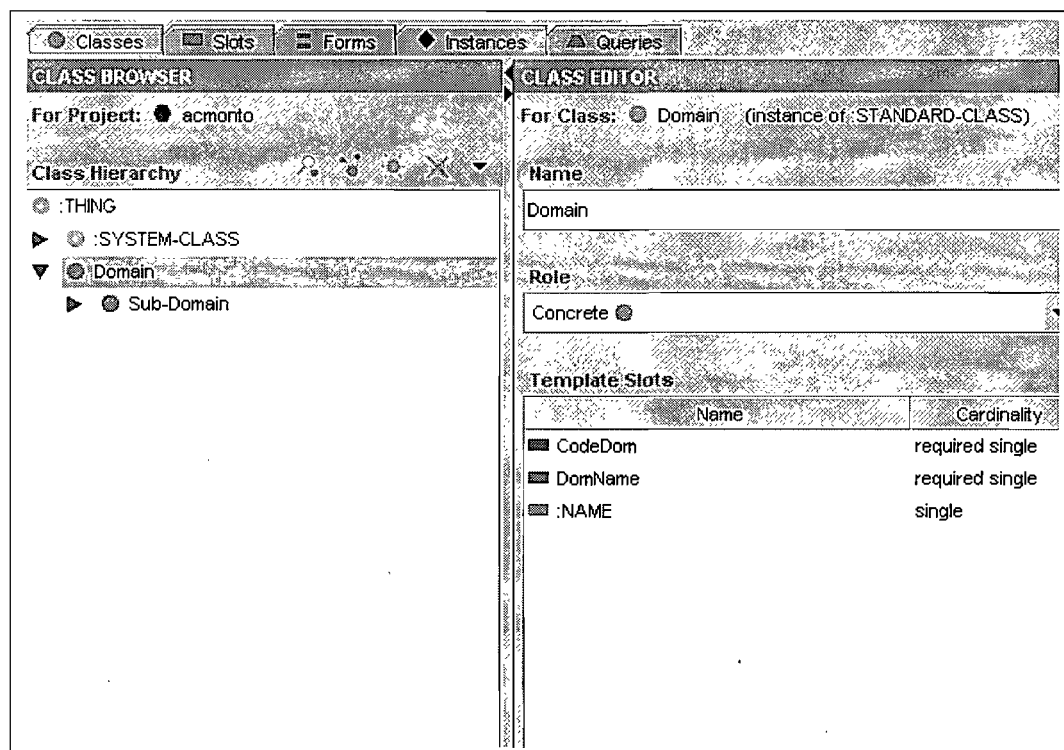


Figure 5.2 – Structure de la classe *Domain* de l'ontologie de LORESA

forme open-source, gratuite qui fournit un ensemble d'outils pour la construction de modèles du domaines ainsi que des applications à base de connaissances basées sur les ontologies. Protégé implémente un ensemble riche de structures de modélisation des connaissances et d'opérations qui permettent la création, la visualisation et la manipulation d'ontologies sous divers formats. Le traitement des ontologies est un domaine très actif où l'on cherche à automatiser au mieux ce processus car il n'existe pas d'approche unique. La construction-enrichissement des ontologies dépend dans un premier lieu des besoins des acteurs ainsi que de l'existant. Pour la construction de notre ontologie nous avons utilisé le langage OWL Lite³(Ontology Web Langage). Ce sous-langage de OWL offre un excellent outil pour la construction d'hierarchies de classification avec des contraintes simples. Ses constructeurs offrent les éléments de base pour la définition de sous-classes et de restrictions de propriétés en plus d'offrir la possibilité de considérer des propriétés

³<http://www.w3.org/TR/2004/REC-owl-features-20040210/>

nécessaires ou optionnelles. Il répond ainsi totalement à notre préoccupation ontologique. Le constructeur de base est la classe *owl:Thing* qui doit être la superclasse de toute classe en OWL. Ainsi nous avons commencé par la définition de la structure de notre hiérarchie traduite par les classes de notre ontologie. Nous avons alors défini la classe *domain* qui sert à organiser notre *connaissance* en sujets. Cette dernière comporte les propriétés suivantes : code du domaine et intitulé du domaine (voir figure 5.2), traduite en OWL par la classe *domain* sous-classe de *owl:Thing* qui comporte deux propriétés : *DomName* et *CodeDom*. L'instanciation de ces classes donne les domaines définis dans la classification de l'ACM (voir figure 5.3).

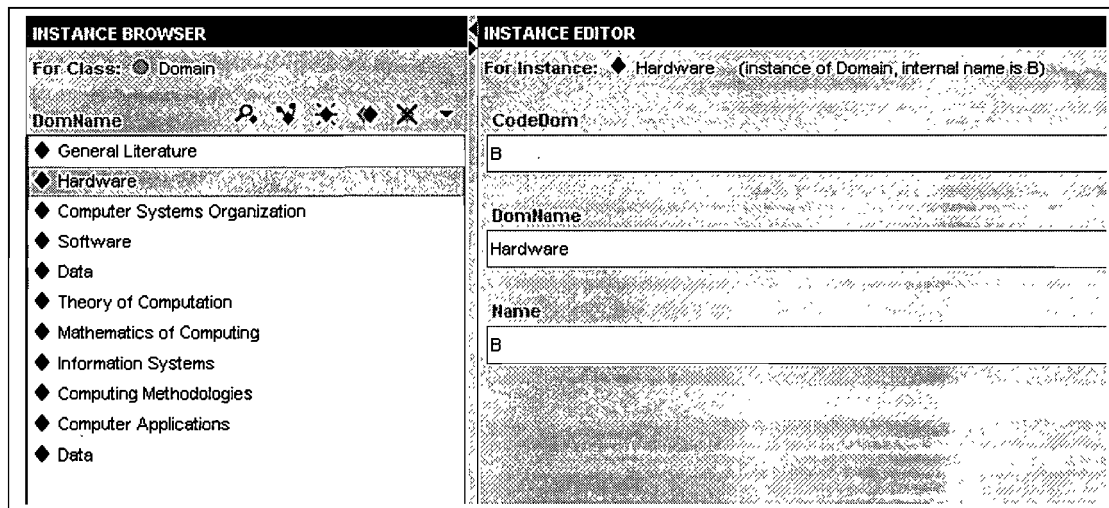


Figure 5.3 – Instances de la classe Domain de l'ontologie de LORESA

Nous avons continué par la définition de la structure de notre ontologie et ce en définissant les sous-classes : *Sub-Domain*, *Category* et *Concept*. (voir figure 5.4)

Dans une première phase d'enrichissement de l'ontologie, nous avons instancié la sous-classe *Sub-Domain* (voir figure 5.5).

Nous avons enchaîné par l'instanciation de la classe *Category* (voir figure 5.6).

Cet enrichissement est basé sur les définitions données dans la classification de l'ACM. Comme nous l'avons spécifié précédemment, pour atteindre un plus haut degré de granularité dans la description des objets d'apprentissage, l'utilisation de l'ontologie inspirée de

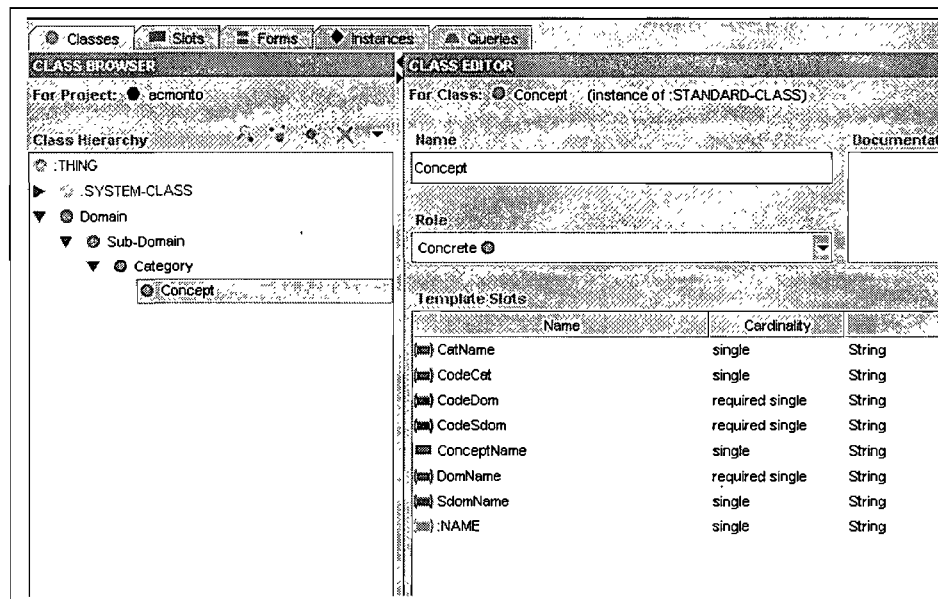


Figure 5.4 – Structure de la hiérarchie des classes de l'ontologie de LORESA

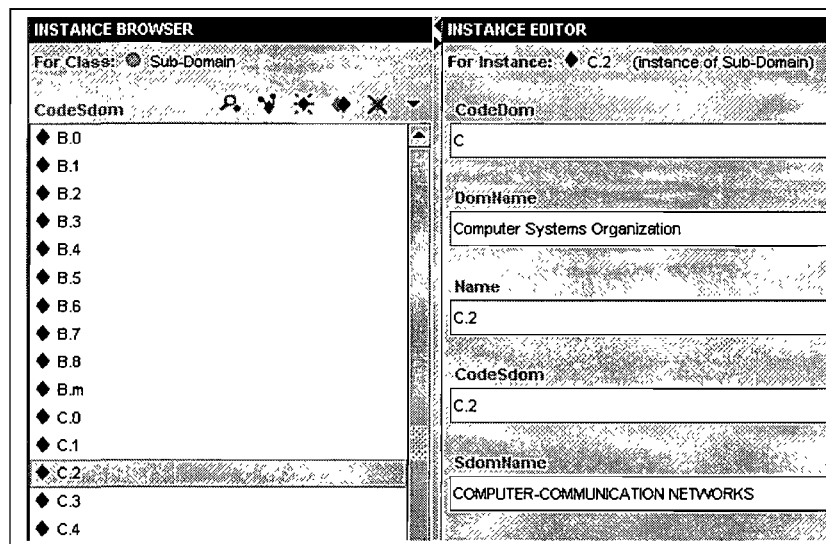


Figure 5.5 – Instance de la classe SubDomain de l'ontologie de LORESA

la classification de L' ACM est insuffisante. À cet effet nous avons défini la classe *Concept* comme sous-classe de *Category*. Son instanciation nous a permis de définir les concepts du cursus informatique (voir figure 5.7). Nous avons ensuite traduit notre ontologie OWL en fichier RDF/XML (voir figure 5.8).

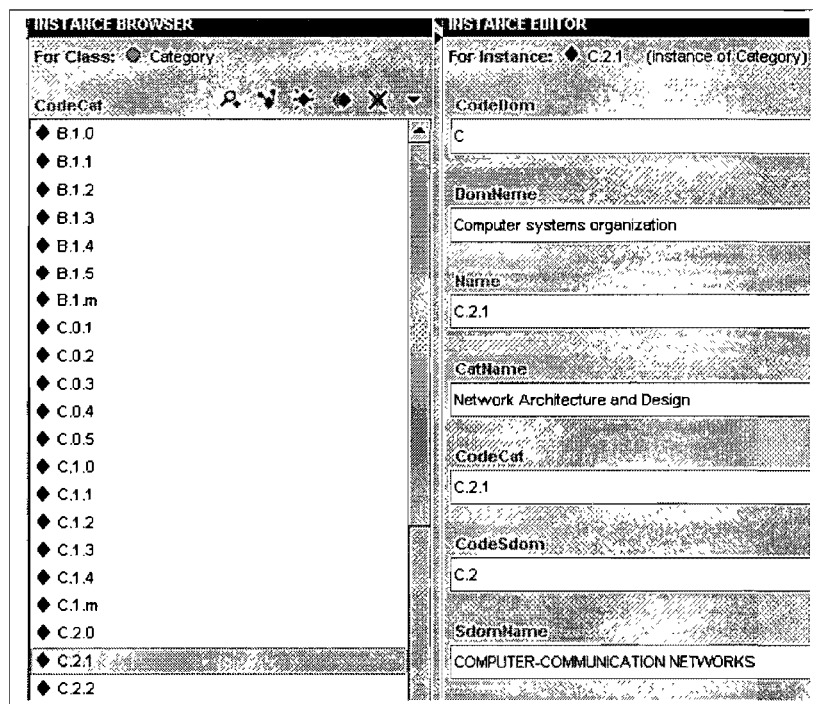


Figure 5.6 – Instance de la classe Category de l'ontologie de LORESA

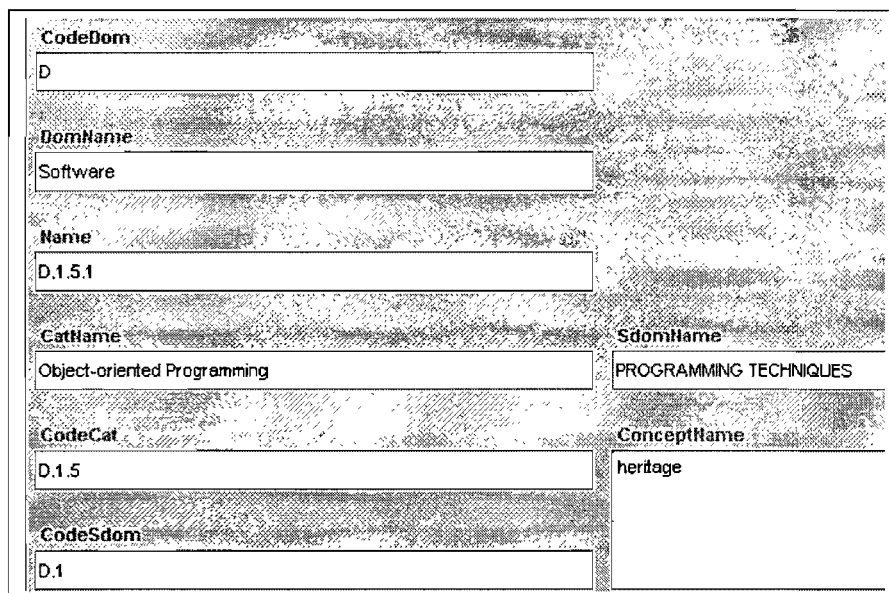


Figure 5.7 – Instance de la classe Concept de l'ontologie de LORESA

```

<Concept rdf:ID="B.1.1.1">
  <DomName rdf:datatype="xsd:string">Hardware</DomName>
  <CodeCat rdf:datatype="xsd:string">B.1.1</CodeCat>
  <SdomName rdf:datatype="xsd:string"
    >CONTROL STRUCTURES AND MICROPROGRAMMING</SdomName>
  <CodeSdom rdf:datatype="xsd:string">B.1</CodeSdom>
  <ConceptName rdf:datatype="xsd:string"
    >Hardwired control</ConceptName>
  <CodeDom rdf:datatype="xsd:string">B</CodeDom>
  <CatName rdf:datatype="xsd:string"
    >Control Design Styles</CatName>
</Concept>
<Concept rdf:ID="B.1.1.2">
  <DomName rdf:datatype="xsd:string">Hardware</DomName>
  <CodeCat rdf:datatype="xsd:string">B.1.1</CodeCat>
  <SdomName rdf:datatype="xsd:string"
    >CONTROL STRUCTURES AND MICROPROGRAMMING</SdomName>
  <CodeSdom rdf:datatype="xsd:string">B.1</CodeSdom>
  <ConceptName rdf:datatype="xsd:string"
    >Microprogrammed logic arrays</ConceptName>
  <CodeDom rdf:datatype="xsd:string">B</CodeDom>
  <CatName rdf:datatype="xsd:string"
    >Control Design Styles</CatName>
</Concept>
<Concept rdf:ID="B.1.1.3">
  <DomName rdf:datatype="xsd:string">Hardware</DomName>
  <CodeCat rdf:datatype="xsd:string">B.1.1</CodeCat>
  <SdomName rdf:datatype="xsd:string"
    >CONTROL STRUCTURES AND MICROPROGRAMMING</SdomName>
  <CodeSdom rdf:datatype="xsd:string">B.1</CodeSdom>
  <ConceptName rdf:datatype="xsd:string"
    >Writable control store</ConceptName>
  <CodeDom rdf:datatype="xsd:string">B</CodeDom>
  <CatName rdf:datatype="xsd:string"
    >Control Design Styles</CatName>
</Concept>

```

Figure 5.8 – Partie du fichier source de l'ontologie de LORESA en RDF/XML

5.3.2 L'entrepôt d'objets d'apprentissage

Nous avons conçu un entrepôt d'objets d'apprentissage qui offre toutes les fonctionnalités décrites en 2.5.1. Les détails de ces fonctionnalités sont donnés dans la section suivante.

5.3.3 Le système d'annotation

Le système d'annotation que nous avons développé dans le cadre de notre travail est fortement inspiré des outils RdfPic et Annotea décrit en 4.3.4.3 et 4.3.4.4 respectivement. Ainsi c'est un système d'annotation sémantique manuel, son principal avantage par rapport à ces derniers réside dans la structure du corps de l'annotation qui n'est pas du

texte libre mais un terme choisi parmi un vocabulaire contrôlé en l'occurrence l'ontologie que nous avons développée et décrit en 5.3.1. Comme nous l'avons spécifié auparavant notre objectif principal par l'annotation de l'objet d'apprentissage est la description de façon formelle des concepts qu'il couvre pour les utiliser ultérieurement à des fins pédagogiques. Pour ce faire, l'auteur d'un objet d'apprentissage qui veut l'annoter est invité par le système à spécifier le sous-domaine et la catégorie auxquels son objet appartient (le domaine étant préalablement spécifié lors de la soumission de l'objet à l'entrepôt). A ce moment, le système charge l'ontologie et grâce à ces informations (sous-domaine et catégorie), il propose la liste des concepts définis sous ladite catégorie.

Nous avons choisi d'annoter l'objet d'apprentissage par le biais de trois concepts : un concept principal et deux concepts secondaires. Ce choix a été fait dans un souci de couverture, en effet il est rare qu'un objet d'apprentissage couvre un seul concept, en fait il traite souvent de plusieurs autres - concept connexes au concept principal-, de manière moins détaillée.

L'annotation se fait donc en choisissant un concept principal et deux concepts secondaires parmi ceux qui sont proposés. L'auteur doit, ensuite, évaluer le degré de leurs couvertures par l'objet en question. La structure de l'annotation est la suivante :

- l'identifiant de l'objet : nous avons choisi d'identifier l'objet à annoter par son adresse URL, car cette information est requise et vérifiée lors de la soumission de l'objet à l'entrepôt d'apprentissage ;
- le sous-domaine : en cherchant dans l'entrepôt l'objet à annoter, le système en extrait alors le domaine. Cette information est utilisée pour charger l'ontologie et plus spécifiquement les sous-domaines définis sous ledit domaine. L'utilisateur en choisit alors un ;
- la catégorie : de la même manière décrite précédemment, en sélectionnant le sous-domaine, le système propose via l'ontologie, un ensemble de catégories qui font partie du sous-domaine choisi ;
- le concept principal : l'utilisateur choisit un concept parmi ceux qui sont proposés

- par l'ontologie ;
- la pertinence pédagogique du concept principal : c'est une valeur numérique prise dans une échelle de 1 à 5, qui indique le degré de couverture pédagogique du concept principal par l'objet. Une pertinence pédagogique de 5 indique une forte couverture pédagogique du contenu de l'objet pour l'acquisition du concept principal ;
 - le concept secondaire 1 : l'utilisateur choisit un concept parmi ceux qui sont proposés par l'ontologie et qui sont différents du concept principal ;
 - la pertinence pédagogique du concept secondaire 1 : c'est une valeur numérique de 1 à 5 (elle doit être inférieure à la pertinence pédagogique au concept principal) ;
 - le concept secondaire 2 : l'utilisateur choisit un concept parmi ceux qui sont proposés par l'ontologie et qui sont différents du concept principal et du concept secondaire 1 ;
 - la pertinence pédagogique du concept secondaire 2 : valeur numérique de 1 à 5. c'est une valeur numérique de 1 à 5 (elle doit être inférieure à la pertinence pédagogique au concept principal) ;

5.3.3.1 Comparaison

Nous présentons dans ce qui suit un tableau (voir tableau 5.1) dans lequel nous comparons notre système d'annotation aux systèmes d'annotation décrits en 4.3.4. Les critères que nous avons choisis sont : les ressources annotées, la méthode d'annotation, les ontologies, les standards, le stockage et l'exploitation des annotations (voir les sections : 4.3.3 et 4.3.4)

Interprétation du tableau

- Les ressources annotées : Les outils : OOF, SMORE et ANNOTEA annotent des ressources dont le contenu est décrit en HTML, XML. Ceci s'explique par le fait que ces outils sont destinés spécifiquement à l'annotation de pages Web. Le champ d'utilisation des outils RdfPic et LORESA est plus restreint puisqu'ils se destinent à

annoter des ressources particulières : les photos JPEG pour le premier et des objets d'apprentissage décrits par des métadonnées LOM pour le deuxième.

- La méthode d'annotation : Tous les outils décrits utilisent une méthode d'annotation manuelle. Ceci revient à choisir manuellement les concepts de l'ontologie qui vont constituer l'annotation. Cette méthode est plus précise que la méthode automatique et est particulièrement adaptée à une utilisation moyenne. Dans le cas où le volume des ressources à annoter est important, elle devient inefficace.

Tableau 5.1 – Tableau comparatif de LORESA avec les outils d’annotation décrits en 5.4

<i>Outils</i>	<i>Ressources annotées</i>	<i>Méthode d’annotation</i>	<i>Ontologies</i>	<i>Standards</i>	<i>Stockage des annotations</i>	<i>Exploitation des annotations</i>
Open Ontology Forge	HTML, texte et image	Manuelle	Ontologies locales éditables	RDF(S), XML, Xlink XPointer et Dublin Core	Local, fichier RDF ou XML	Agents logiciels
SMORE	HTML, texte, image et E-Mail	Manuelle	Serveur d’ontologies éditables	RDF(S)	Local, intégration à la page	Agents logiciels
RdfPic	Image JPEG	Manuelle	NA : Non Appllicable	RDF (S), Dublin Core	Local, intégration à l’image	Recherche d’images
Annotea	HTML, XHTML et XML	Manuelle	NA : Non Appllicable	RDF(S) XLink, XPointer	Serveur d’annotation	Classification automatique, moteurs de recherche
LORESA	Objet d’apprentissage	Manuelle	Ontologie consultable	OWL, RDF/XML	Serveur d’annotation	Recommandation sémantique d’objets d’apprentissage

- les ontologies : tous les outils à l’exception de RdfPic et ANNOTEA, utilisent des ontologies. Les outils OOF et SMORE ont opté pour l’utilisation d’ontologies locales et éditables. Dans un souci de contrôle de la cohérence, nous avons opté pour l’ontologie de LORESA d’offrir la possibilité de la consultation pour tous les usagers. Nous avons réservé la possibilité de l’édition pour les utilisateurs dont le rôle est : *expert* ;
- les standards : tous les outils utilisent le langage de définition d’ontologies : RDFS. LORESA est le seul outil à utiliser le langage OWL. Cet outil est de plus en plus utilisé pour la définition des ontologies sur le Web ;
- le stockage des annotations : seuls les outils ANOTEA et LORESA stockent les annotations réalisées dans un serveur d’annotation. Ceci permet à n’importe quel utilisateur de pouvoir consulter les annotations d’une ressource ce qui n’est pas possible dans le cas où les annotations sont stockées localement (cas des autres outils) où elles ne peuvent être consultées que par l’utilisateur qui les a créées ;
- l’exploitation des annotations : tous les outils exploitent les annotations créées pour améliorer l’utilisation des ressources qu’elles annotent. Le système LORESA utilise les annotations sémantiques pour l’aide à la recherche des objets d’apprentissage. Il permet particulièrement, grâce à ces annotations, d’indexer ces ressources et d’aider à les recommander à des utilisateurs ;

5.3.4 Le système de recommandation

Le point le plus important dans la conception des systèmes de recommandation est sans aucun doute le choix des informations qui servent à décrire les ressources à recommander. Ces données sont porteuses de *la sémantique* de la ressource et de sa *valeur*. Selon [Resnick et *al.*, 1997], il est primordial de prendre en compte plusieurs éléments lors de la conception d’un système de recommandation basé sur le filtrage collaboratif parmi lesquels : l’ontologie de la structure des métadonnées, les métadonnées utilisées et celles qui sont utilisées pour l’évaluation. les standards tels que LOM, se sont concentrés

sur les aspects techniques de la description des objets d'apprentissage dans un entrepôt de ressources. Ils ne s'intéressent pas à la description de la couverture pédagogique de ces objets du contenu qu'ils traitent.

5.3.4.1 L'évaluation de la pertinence pédagogique

C'est en réponse à la lacune de la description de la couverture pédagogique des objets d'apprentissage et respectant les impératifs de conception de systèmes de recommandation, que nous avons conçu LORESA. Le système de recommandation ou le recommandeur est le cœur de notre architecture. Il sert à recommander des objets d'apprentissage en se basant dans un premier temps sur les préférences de l'utilisateur et en prenant en compte ensuite son profil pour lui recommander les ressources les plus adéquates à sa requête. Notre contribution majeure se situe, en fait, dans la définition de nouvelles métadonnées - les annotations sémantiques- pour pouvoir décrire l'objectif pédagogique de l'objet en question de façon formelle et partagée et d'y appliquer les techniques du filtrage collaboratif. Comme nous l'avons vu en 5.3.3, l'objet est annoté une seule fois par son auteur lors de sa soumission à l'entrepôt et cette annotation est présentée comme recommandation initiale de l'objet (lorsque l'objet n'a pas encore été évalué par les autres usagers). Cependant n'importe quel autre utilisateur peut évaluer ces métadonnées dotant ainsi l'objet de plusieurs évaluations qui contribuent à la personnalisation des recommandations produites. Lors de la création de descriptions d'objets d'apprentissage, il est important de décrire de manière précise dans quel contexte on peut utiliser cet objet. Les champs sujets, mots-clés décrits dans le standard LOM, même s'ils aident à donner une idée de ce contexte, ceci reste cependant ambigu et sujet à confusion. C'est dans ce but que nous avons proposé des métadonnées additionnelles qui servent à décrire ce contexte d'utilisation de façon non-ambigüe. Ainsi comme nous avons vu en 6.3.2, la structure des métadonnées que nous avons ajoutées sous forme d'annotation sémantique comportent entre autres : trois concepts qui servent à déterminer de façon formelle le sujet couvert et surtout le degré de cette couverture. Ce sont donc ces concepts conjugués à leurs degrés

de couverture que nous utilisons dans notre technique de filtrage collaboratif. Ainsi le vote d'évaluation d'un utilisateur i de l'objet d'apprentissage j est un vecteur de dimension 3 noté par :

$$\begin{pmatrix} CPr_1(i, j) \\ CPr_2(i, j) \\ CPr_3(i, j) \end{pmatrix}$$

$CPr_1(i, j)$ représente la pertinence pédagogique du concept principal de l'objet d'apprentissage j évaluée par l'utilisateur i . $CPr_2(i, j)$, $CPr_3(i, j)$ représentent les pertinences pédagogiques des concepts secondaires un et deux respectivement.

Nous utilisons dans notre système l'algorithme de filtrage collaboratif basé mémoire (voir 3.2.1.1) où nous calculons les prédictions des évaluations des pertinences pédagogiques d'un utilisateur actif a pour l'objet j par le biais de la formule suivante :

$$\begin{pmatrix} CPr_1(a, j) \\ CPr_2(a, j) \\ CPr_3(a, j) \end{pmatrix} = \begin{pmatrix} \overline{CPr_1}(a) \\ \overline{CPr_2}(a) \\ \overline{CPr_3}(a) \end{pmatrix} + \frac{\sum_{i=1}^n ws(a, i) \times \begin{pmatrix} CPr_1(i, j) - \overline{CPr_1}(i) \\ CPr_2(i, j) - \overline{CPr_2}(i) \\ CPr_3(i, j) - \overline{CPr_3}(i) \end{pmatrix}}{\sum_{i=1}^n |w(a, i)|} \quad (5.1)$$

Les vecteurs $\begin{pmatrix} \overline{CPr_1}(a) \\ \overline{CPr_2}(a) \\ \overline{CPr_3}(a) \end{pmatrix}$ et $\begin{pmatrix} \overline{CPr_1}(i) \\ \overline{CPr_2}(i) \\ \overline{CPr_3}(i) \end{pmatrix}$ sont les votes moyens des utilisateurs a et i respectivement où $\overline{CPr_1}(a)$ représente la moyenne des évaluations de l'utilisateur a pour la pertinence pédagogique du concept principal pour tous les objets qu'il a évalués. $\overline{CPr_2}(a)$ et $\overline{CPr_3}(a)$ sont les moyennes des évaluations de l'utilisateur a pour les pertinence pédagogiques des deux concepts secondaires pour tous les objets qu'il a évalués. Il en est de même pour $\overline{CPr_1}(i)$, $\overline{CPr_2}(i)$ et $\overline{CPr_3}(i)$ qui concernent les moyennes des évaluations de l'utilisateur i . La valeur de n représente le nombre des utilisateurs dans la base utilisateurs/votes qui ont évalué l'objet j . Quant à la valeur de $ws(a, i)$, elle désigne

la similarité entre chaque utilisateur i et l'utilisateur actif a .

5.3.4.2 La similarité

Pour le calcul de la similarité entre deux utilisateurs de LORESA, nous nous basons sur les évaluations communes d'objets entre ces deux utilisateurs. Nous avons commencé par définir : *la pertinence pédagogique* : Pr comme une moyenne pondérée des évaluations des couvertures pédagogiques des concepts qui annotent l'objet d'apprentissage. Elle représente sa profondeur pédagogique de façon générale. Nous avons alors accordé la pondération de 3 à l'évaluation de la couverture du concept principal qui doit détenir la sémantique principale de l'objet. Les deux autres concepts ont chacun la même importance dans le calcul de la pertinence pédagogique de l'objet traduite par une pondération égale à l'unité. Cette mesure est donnée par la formule suivante :

$$Pr_{i,j} = \frac{1}{5} \times (3 \times CPr_1(i, j)) + CPr_2(i, j) + CPr_3(i, j) \quad (5.2)$$

$Pr_{i,j}$ représente la pertinence pédagogique de l'objet d'apprentissage j calculée à partir de l'évaluation de l'utilisateur i des pertinence pédagogiques des concepts annotants. Ensuite nous appliqué la formule du calcul de la similarité vectorielle donnée par la formule 3.3 dans la section 3.2.1.1. qui donne le similarité entre l'utilisateur courant a et un utilisateur j par le biais de la formule suivante :

$$ws(a, i) = \frac{\sum_{j=1}^m pr_{a,j} \times pr_{i,j}}{\sqrt{\sum_j (pr_{a,j})^2 \times \sum_j (pr_{i,j})^2}} \quad (5.3)$$

où :

- m est le nombre des items évalués par i et a .
- $pr_{a,j}$ est l'évaluation de l'utilisateur a de la pertinence pédagogique de l'objet j .
- $pr_{i,j}$ est l'évaluation de l'utilisateur i de la pertinence pédagogique de l'objet j .

Pour le calcul de la prédiction de l'évaluation de l'objet i pour l'utilisateur a , il est nécessaire de déterminer le voisinage de ce dernier. Pour y arriver, nous avons effectué des calculs préliminaires qui nous ont aidé à définir le voisinage d'un utilisateur a comme étant tous les usagers i tel que : $ws(a,i) \geq seuil$. Dans le chapitre implémentation nous allons définir la valeur de la constante *seuil*.

5.3.4.3 Génération de l'explication

La plupart des systèmes de recommandations présentent leurs recommandations aux utilisateurs sans leur fournir l'explication du processus qui a conduit à leur génération. Nous avons voulu combler cette lacune en offrant à l'utilisateur une connaissance supplémentaire qui peut l'aider dans son processus de choix. Cette opération est réalisée par le système à la demande de l'utilisateur. Ce dernier peut, pour une recommandation donnée obtenir le processus qui a conduit à sa génération et qui peut prendre trois formes :

- soit que la recommandation est l'annotation qui a été réalisée par son auteur (objet non-évalué par les autres utilisateurs) ;
- soit que la recommandation est générée par les évaluations des utilisateurs similaires au demandeur. Dans ce cas les pseudonymes de ces utilisateurs ainsi que leurs évaluations lui sont présentés ;
- soit que la recommandation est l'évaluation du demandeur lui-même.

5.4 Architecture de LORESA

Dans cette section, nous présentons l'architecture générale LORESA (voir Figure 6.9) sous forme de plusieurs composantes qui dans leurs diverses interactions permettent d'offrir de façon transparente les fonctionnalités de : l'entrepôt d'objet d'apprentissage, le système d'annotation et le système de recommandation. Ainsi nous allons d'abord commencer par les modules qui réalisent les fonctionnalités de l'entrepôt d'objet d'apprentissage : module d'ajout d'objets d'apprentissage et le module de recherche d'objets d'apprentis-

sage. Nous enchaînons ensuite par les modules qui réalisent la fonctionnalité de l'annotation sémantique : module d'annotation sémantique, module de consultation de l'ontologie et le module d'enrichissement de l'ontologie suit enfin les modules qui réalisent la fonctionnalité de la recommandation : le module de gestion des profils utilisateurs, le module de la recommandation sémantique et le module de l'évaluation de la pertinence pédagogique. Nous allons commencer par module de gestion du profil de l'utilisateur, module faisant partie du sous-système de recommandation et qui représente le module central du système puisqu'il permet de gérer toutes les opérations qu'effectuent les utilisateurs pour permettre de générer les recommandations.

5.4.1 Composant de gestion de profil

C'est le module qui représente le point d'entrée à notre système. Il assure les fonctions suivantes :

- *La création des profils* des utilisateurs lors de la phase de l'enregistrement (voir figure 5.9, flèches étiquetées 0). Ils doivent remplir un formulaire avec leurs données personnelles : pseudonyme, mot de passe, courriel ainsi que d'autres données concernant leurs préférences relatives aux ressources recherchées. La phase d'enregistrement est effectuée une seule fois pour chaque nouvel utilisateur.
- *L'authentification des utilisateurs* lors de leurs connexions ultérieures au système.
- *La mise à jour des profils des utilisateurs déjà enregistrés* : il ajoute les objets qu'ils ont soumis à l'entrepôt et les objets qu'ils ont évalués (voir figure 5.9, flèches étiquetées 1,2 et 8 respectivement), les changements dans leurs préférences, etc.
- *La recherche des profils similaires*, lors du calcul d'une recommandation (voir figure 5.9, flèche étiquetée 5).

5.4.2 Base des profils

Comme nous l'avons vu précédemment, le module de gestion des profils doit constamment mettre à jour le profil utilisateur. En effet, le recommandeur d'objets se base sur les

dynamiques.

1. *Les données statiques :*

Ces informations sont recueillies lors de l'enregistrement de l'utilisateur et servent à lui faire les premières recommandations qui se basent sur le filtrage à base de contenu. Ces informations ne changent pas d'une session à l'autre, à moins que l'utilisateur veuille changer ses préférences via le composant de gestion du profil. Citons parmi ces informations le pseudonyme, que l'usager utilise pour s'identifier, une adresse courriel valide, la combinaison de ces deux informations permet de garantir l'unicité de l'usager. Un mot de passe d'une longueur minimale de quatre caractères, le rôle de l'usager qui peut être : apprenant ou enseignant, un champ d'intérêt qui décrit le domaine dans lequel l'usager pourrait chercher les objets d'apprentissage qui l'intéressent, le format et la langue préférés des ressources recherchées. Ces informations sont utilisées pour présenter à l'utilisateur à chaque fois qu'il se connecte au système, des ressources présentes dans l'entrepôt et qui peuvent potentiellement l'intéresser. Ce sont les recommandations basées sur le filtrage à base de contenu. Les données statiques de l'usager peuvent être représentées par le vecteur suivant : (*pseudo, email, pass, role, inter, form, lang*)

2. *Les données dynamiques :*

Ces données changent au fur et à mesure que l'usager utilise LORESA. Ainsi, à chaque fois qu'il soumet un objet à l'entrepôt, l'annote et évalue les recommandations présentées, le composant de gestion de profil enregistre les données correspondantes dans son profil. Toutes ces informations servent à mettre à jour le voisinage de l'utilisateur et d'améliorer ainsi les recommandations qui lui sont présentées. Les données dynamiques de l'usager peuvent être représentées par le vecteur suivant : (*Objets-soumis, Objets-annotés, Objets-évalués*).

5.4.3 Composant d'ajout d'objets d'apprentissage

C'est le module qui permet à un utilisateur enregistré de soumettre un objet d'apprentissage à l'entrepôt des ressources pédagogiques (voir figure 5.9, flèche étiquetée 1). Il permet donc d'ajouter à l'entrepôt une ressource pédagogique définie par un identifiant unique qui est son adresse URL et décrite par des métadonnées faisant partie du modèle LOM parmi lesquelles : titre, format, auteur, domaine , difficulté relative, audience cible, etc.

5.4.4 Composant de recherche d'objets d'apprentissage

Ce composant permet à un utilisateur enregistré, de rechercher parmi la banque des ressources celles qui répondent à des critères spécifiques (voir figure 5.9, flèche étiquetée 3) . Pour ce faire l'utilisateur doit introduire via la page de la requête les critères recherchés. Il renseigne à cet effet, les champs des métadonnées spécifiques dans la page de la requête. Le système présente alors à l'utilisateur, les objets d'apprentissages qui répondent aux critères introduits. À ce niveau, LORESA, offre une nouvelle fonctionnalité à l'usager en l'occurrence : la recherche sémantique. Celle-ci permet, d'afficher pour chaque ressource - lorsqu'elle a été annotée (voir 5.4.5)- , les concepts qu'elle couvre.

5.4.5 Composant d'annotation d'objets d'apprentissage

C'est le module qui permet à un utilisateur enregistré et qui a soumis un objet d'apprentissage à l'entrepôt, de l'annoter par le biais de l'ontologie (voir figure 5.9, flèche étiquetée 2) . Comme nous l'avons vu en 5.3.3, lors de l'enregistrement d'un objet d'apprentissage dans l'entrepôt, l'utilisateur est invité à choisir le domaine du sujet qu'il traite. Cette information est utilisée par le système d'annotation pour charger à partir de l'ontologie le domaine adéquat. L'usager est alors invité à choisir dans cet ordre : le sous-domaine, la catégorie et trois concepts parmi ceux qui existent dans l'ontologie. En choisissant le concept principal et les deux concepts secondaires, l'annotateur est appelé à donner des

évaluations numériques initiales, qui décrivent le degré de contribution de l'objet à l'acquisition de chaque concept choisi. Nous avons appelé chaque valeur : pertinence pédagogique du concept. Le sous-domaine, la catégorie, les 03 concepts évalués constituent l'annotation sémantique de cet objet d'apprentissage.

5.4.6 Composant d'enrichissement de l'ontologie

C'est le module qui permet à un utilisateur qui dispose des privilèges requis (expert), d'ajouter des concepts au niveau de l'ontologie (voir figure 5.9, flèche étiquetée 9). Notre démarche ontologique consiste à étendre la classification de l'ACM du cursus informatique, pour qu'elle puisse avoir un degré plus élevé de granularité (voir 5.3.1). Nous avons opté pour cette fonctionnalité de ne permettre à un utilisateur expert, que l'**instanciation** de la **classe concept**. Ainsi nous garantissons la cohérence et la stabilité de notre ontologie.

5.4.7 Composant de consultation de l'ontologie

C'est le module qui permet à tous les usagers de consulter l'ontologie utilisée pour l'annotation des objets d'apprentissage. (voir figure 5.9, flèche étiquetée 10). Notre démarche ontologique consiste à étendre la classification de l'ACM du cursus informatique, pour qu'elle puisse avoir un degré plus élevé de granularité (voir 6.3.1). Nous avons opté pour cette fonctionnalité de ne permettre à un utilisateur expert, que l'**instanciation** de la **classe concept**. Ainsi nous garantissons la cohérence et la stabilité de notre ontologie.

5.4.8 Composant de recommandation des objets d'apprentissage

C'est le cœur de notre architecture puisque son rôle est de fournir à la demande d'un usager, des recommandations d'objets d'apprentissage, qui pourraient répondre au mieux à sa requête. Pour cela, il se base d'abord sur les données enregistrées dans le profil de l'apprenant lors de son enregistrement ensuite, sur les données des autres utilisateurs qui

lui ressemblent (voir figure 5.9, flèches 5 et 6). Notre recommandeur d'objets d'apprentissage s'appuie sur une technique de recommandation hybride. Elle est basée sur la mise en cascade des deux techniques de filtrage : le filtrage à base de contenu et le filtrage collaboratif. La première technique est appliquée en premier, ensuite, et à la demande de l'utilisateur, ses résultats sont utilisés comme entrée à une technique de filtrage collaboratif.

5.4.9 Composant d'évaluation des pertinences pédagogiques

Ce module permet à l'utilisateur de donner son feedback concernant les recommandations qu'il a reçues. Il est alors invité à donner les valeurs, qu'il croit, représentent mieux les pertinences pédagogiques des concepts pour l'objet qu'il évalue (voir figure 5.9, flèche étiquetée 7) . Ces nouvelles valeurs vont être enregistrées dans le profil de l'utilisateur et vont servir à calculer les recommandations pour les usagers similaires à lui. Nous présentons ci-après le pseudo-code du recommandeur :

- 1: **si** l'utilisateur u enregistré fait une requête de recherche d'objets **alors** Afficher les objets correspondants à ses préférences enregistrées dans son profil
- 2: **si** l'utilisateur veut obtenir des recommandations concernant les objets sélectionnés **alors**
- 3: **pour** chaque objet sélectionné **faire**
- 4: **si** l'objet n'a pas été évalué par d'autres usagers **alors** Afficher les pertinences pédagogiques initiales données par l'annotateur de l'objet
- 5: **sinon**
 - Calculer la similarité de u avec les autres usagers (voir Formule 6.3).
 - Déterminer le voisinage de u .
 - Calculer les prédictions de pertinence pédagogique correspondante à chaque concept (voir formule 6.1)
- 6: **fin si**
- 7: **si** l'utilisateur veut connaître le processus de recommandation **alors** Afficher les évaluations de ses voisins et/ou les évaluations initiales

```

8:          fin si
9:          fin pour
10:        fin si
11:        si l'utilisateur u est satisfait des recommandations alors Enregistrer ces valeurs
           dans le profil utilisateur
12:        sinon
           – Demander l'évaluation des recommandations
           – Enregistrer les valeurs dans le profil de l'usager
13:        fin si

```

Pour mieux comprendre le fonctionnement de l'algorithme du recommandeur de LORESA, nous présentons un scénario d'utilisation complet. Soit Alice une utilisatrice déjà enregistrée dans LORESA. Le vecteur qui représente les données statiques de son profil est donné par : (*Alice*, *alice@yahoo.com*, ******, *Apprenant*, *Computer Systems*, *PDF*, *Anglais*), quant aux données dynamiques de son profil, elles sont données par le vecteur : (*(objet6)*, (*objet6*), (*objet1*, *objet2*, *objet3*)). Alice se connecte à LORESA, le système lui recommande une liste de ressources d'objets qui répondent aux critères spécifiés dans la partie statique de son profil : des ressources du domaine *Computer Systems*, de format *PDF* et en langue *anglaise*. Cette liste est donnée par : (*objet4*, *objet5*). Alice est intéressée par l'objet 4 et voudrait avoir les recommandations des autres utilisateurs. Pour cela nous avons besoin des évaluations des autres utilisateurs. Prenons quatre utilisateurs de LORESA qui ont chacun évalué quatre objets au maximum. Nous avons ainsi les évaluations de chacun pour les concepts de chaque objet. Nous regroupons ces votes dans une matrice (*utilisateurs* \times *objets*), ensuite nous appliquons la méthode du *cosinus* qui consiste à définir le voisinage d'un utilisateur donné, et nous passons à la génération des recommandations en se basant sur son voisinage.

Le Tableau 6.2 est un exemple qui représente une matrice des évaluations-quand elles existent- de cinq utilisateurs pour les trois concepts de chaque objet des quatre objets évalués. Chaque objet a trois évaluations qui correspondent aux pertinences pédagogiques

Tableau 5.2 – Tableau des évaluations des utilisateurs de LORESA pour les quatre objets d'apprentissage

	<i>Objet1</i>	<i>Objet2</i>	<i>Objet3</i>	<i>Objet4</i>
Alice	4	3	2	X
	1	1	1	Y
	2	2	1	Z
Éve	3	2	3	3
	2	1	2	1
	2	1	2	2
Fred	2	1	2	4
	1	1	1	1
	1	1	1	1
Jean	4	3	1	4
	1	1	1	2
	2	1	1	2
Dave	1	NE	5	2
	1	NE	4	1
	1	NE	4	1

des concepts annotants (principal et secondaires). Les valeurs des évaluations varient de 1 à 5. Quand un utilisateur n'a pas encore évalué un objet, - le cas de Dave pour l'objet *Objet2* - l'évaluation correspondante est notée par : NE (Not Evaluated). Les autres valeurs correspondent aux évaluations données par l'utilisateur aux pertinences pédagogiques de chaque concept. Nous voulons par cet exemple, générer les prédictions des pertinences pédagogiques de *l'objet4* pour Alice. Nous cherchons donc à prévoir, les votes que pourrait donner Alice à chacune des pertinences pédagogiques des concepts annotants. Nous nous basons pour calculer ces prédictions, sur son voisinage qui est déterminé à l'aide des formules 6.2 et 6.3. Dans le Tableau 6.2, les valeurs : **X**, **Y** et **Z** sont les prédictions que nous cherchons à calculer. Nous allons suivre les étapes suivantes :

1. *Calculer la matrice des Pertinences pédagogiques des objets :*

Avant de pouvoir calculer les similarités entre Alice et les différents utilisateurs, nous calculons d'abord la matrice des évaluations des pertinences pédagogiques des objets selon la formule : 6.2, ainsi nous obtenons le tableau suivant :

Tableau 5.3 – Tableau des Pertinences pédagogiques des quatre objets d'apprentissage

	$Pr_{User,1}$	$Pr_{User,2}$	$Pr_{User,2}$	$Pr_{User,2}$
Alice	3	2,2	1,6	Pr_{Ob4}
Éve	2,6	1,6	2,6	2,4
Fred	1,6	1	1,6	2,8
Jean	3	2,2	1	3,2
Dave	1	?	4,6	1,6

ainsi la valeur de $Pr_{Alice,1}$ est donnée par :

$$Pr_{Alice,1} = \frac{(3 \times 4) + 1 + 2}{5} = 3$$

Il en est de même pour les valeurs $Pr_{Alice,2}$ et $Pr_{Alice,3}$ données respectivement par

$$Pr_{Alice,2} = \frac{(3 \times 3) + 1 + 1}{5} = 2,6$$

$$Pr_{Alice,3} = \frac{(2 \times 3) + 1 + 1}{5} = 1,6$$

Les valeurs $Pr_{Eve,1}$, $Pr_{Eve,2}$, $Pr_{Eve,3}$ et $Pr_{Eve,4}$ sont données par :

$$Pr_{Eve,1} = \frac{(3 \times 3) + 2 + 2}{5} = 2,6$$

$$Pr_{Eve,2} = \frac{(3 \times 2) + 1 + 2}{5} = 1,6$$

$$Pr_{Eve,3} = \frac{(3 \times 3) + 2 + 2}{5} = 2,6$$

$$Pr_{Eve,4} = \frac{(3 \times 3) + 1 + 2}{5} = 2,4$$

Les valeurs $Pr_{Fred,1}$, $Pr_{Fred,2}$, $Pr_{Fred,3}$ et $Pr_{Fred,4}$ sont données par :

$$Pr_{Fred,1} = \frac{(3 \times 2) + 1 + 1}{5} = 1,6$$

$$Pr_{Fred,2} = \frac{(3 \times 1) + 1 + 1}{5} = 1$$

$$Pr_{Fred,3} = \frac{(3 \times 2) + 1 + 1}{5} = 1,6$$

$$Pr_{Fred,4} = \frac{(3 \times 4) + 1 + 1}{5} = 2,8$$

Les valeurs $Pr_{Jean,1}$, $Pr_{Jean,2}$, $Pr_{Jean,3}$ et $Pr_{Jean,4}$ sont données par :

$$Pr_{Jean,1} = \frac{(3 \times 2) + 1 + 1}{5} = 1,6$$

$$Pr_{Jean,2} = \frac{(3 \times 1) + 1 + 1}{5} = 1$$

$$Pr_{Jean,3} = \frac{(3 \times 2) + 1 + 1}{5} = 1,6$$

$$Pr_{Jean,4} = \frac{(3 \times 4) + 1 + 1}{5} = 2,8$$

Les valeurs $Pr_{Dave,1}$ et $Pr_{Dave,4}$ sont données par :

$$Pr_{Dave,1} = \frac{(3 \times 1) + 1 + 1}{5} = 1$$

$$Pr_{Dave,3} = \frac{(3 \times 5) + 4 + 4}{5} = 4,6$$

$$Pr_{Dave,4} = \frac{(3 \times 2) + 1 + 1}{5} = 1,6$$

2. *Calculer les similarités entre Alice et les autres usagers :*

Nous utilisons à cet effet la formule 6.3 ainsi nous obtenons :

$$ws(Alice, Eve) = \frac{(3 * 2,6) + (2,2 * 1,6) + (1,6 * 2,6)}{\sqrt{((3)^2 + (2,2)^2 + (1,6)^2) * ((2,6)^2 + (1,6)^2 + (2,6)^2)}} = 0,950$$

$$ws(Alice, Eve) = 0,95$$

$$ws(Alice, Fred) = \frac{(3 * 1,6) + (2,2 * 1) + (1,6 * 1,6)}{\sqrt{((3)^2 + (2,2)^2 + (1,6)^2) * ((1,6)^2 + (1)^2 + (1,6)^2)}} = 0,956$$

$$ws(Alice, Fred) = 0,96$$

$$ws(Alice, Jean) = \frac{(3 * 2,6) + (2,2 * 1,6) + (1,6 * 2,6)}{\sqrt{((3)^2 + (2,2)^2 + (1,6)^2) * ((2,6)^2 + (1,6)^2 + (2,6)^2)}} = 0,988$$

$$ws(Alice, Jean) = 0,99$$

$$ws(Alice, Dave) = \frac{(3 * 1) + (1,6 * 4,6)}{\sqrt{((3)^2 + (1,6)^2) * ((1)^2 + (4,6)^2)}} = 0,64$$

$$ws(Alice, Dave) = 0,64$$

3. Déterminer le voisinage d'Alice :

Nous pouvons maintenant sélectionner les utilisateurs qui possèdent un coefficient de similarité supérieur à un certain seuil k . Nous avons choisi de prendre $k = 0,75$ ce qui amène à retenir Éve, Fred et Jean. Nous éliminons Dave car sa similarité avec Alice est inférieur à $0,75$. Ces trois utilisateurs constituent le voisinage d'Alice,

le choix de k est fait de telle sorte à maximiser la précision des recommandations calculées.

4. Générer les recommandations :

Afin de calculer les différentes prédictions d'Alice pour chaque pertinence pédagogique de l'objet *Objet4*, nous utilisons la Formule 6.1. Nous avons donc besoin de calculer les évaluations moyennes d'Alice ainsi que celles de tous les usagers similaires à Alice :

$$\begin{pmatrix} \overline{CPr_1(Alice)} \\ \overline{CPr_2(Alice)} \\ \overline{CPr_3(Alice)} \end{pmatrix} = \frac{1}{3} \times \begin{pmatrix} 4+3+2 \\ 1+1+1 \\ 2+2+1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1,66 \end{pmatrix}$$

De la même façon nous calculons le vecteur de la moyennes des évaluations d'Éve :

$$\begin{pmatrix} \overline{CPr_1(Eve)} \\ \overline{CPr_2(Eve)} \\ \overline{CPr_3(Eve)} \end{pmatrix} = \frac{1}{4} \times \begin{pmatrix} 3+2+3+3 \\ 2+1+2+1 \\ 2+1+2+2 \end{pmatrix} = \begin{pmatrix} 2,75 \\ 1,5 \\ 1,75 \end{pmatrix}$$

de Fred :

$$\begin{pmatrix} \overline{CPr_1(Fred)} \\ \overline{CPr_2(Fred)} \\ \overline{CPr_3(Fred)} \end{pmatrix} = \frac{1}{4} \times \begin{pmatrix} 2+1+2+4 \\ 1+1+1+1 \\ 1+1+1+1 \end{pmatrix} = \begin{pmatrix} 2,25 \\ 1 \\ 1 \end{pmatrix}$$

et de Jean :

$$\begin{pmatrix} \overline{CPr_1(Jean)} \\ \overline{CPr_2(Jean)} \\ \overline{CPr_3(Jean)} \end{pmatrix} = \frac{1}{4} \times \begin{pmatrix} 4+3+1+4 \\ 1+1+1+2 \\ 2+1+1+2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1,25 \\ 1,5 \end{pmatrix}$$

Nous pouvons maintenant générer les recommandations pour Alice en calculant

d'abord :

$$\sum_{i=1}^3 |w(Alice, i)| = 0,95 + 0,96 + 0,99$$

$$\sum_{i=1}^3 |w(Alice, i)| = 2,9$$

ensuite nous générons les recommandations :

$$\begin{pmatrix} CPr_1(Alice, Objet4) \\ CPr_2(Alice, Objet4) \\ CPr_3(Alice, Objet4) \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1,66 \end{pmatrix} + \frac{0,95}{2,9} \begin{pmatrix} 3-2,75 \\ 1-1,5 \\ 2-1,5 \end{pmatrix} + \frac{0,96}{2,9} \begin{pmatrix} 4-2,25 \\ 1-1 \\ 1-1 \end{pmatrix} + \frac{0,99}{2,9} \begin{pmatrix} 4-3 \\ 2-1 \\ 2-1,5 \end{pmatrix}$$

d'où :

$$\begin{pmatrix} CPr_1(Alice, Objet4) \\ CPr_2(Alice, Objet4) \\ CPr_3(Alice, Objet4) \end{pmatrix} = \begin{pmatrix} 4,07 \\ 1,12 \\ 2,37 \end{pmatrix}$$

$$\begin{pmatrix} CPr_1(Alice, Objet4) \\ CPr_2(Alice, Objet4) \\ CPr_3(Alice, Objet4) \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}$$

5. Résultats et Feedback :

Le système présente à Alice les prédictions des pertinences pédagogiques pour l'objet d'apprentissage : *Objet4*, ainsi que pour tous les objets qu'elle aurait sélectionnés. Elle peut obtenir à la demande, des explications au sujet des recommandations présentées par le système. Nous avons voulu par le biais de cette option, donner à l'utilisateur un moyen supplémentaire d'aide au choix en lui expliquant *quelles informations* ont été prises en compte (ses voisins et leurs évaluations) pour le *calcul des recommandations générées*. Ensuite, Alice peut consulter l'objet proposé ce qui lui permet de juger des pertinences pédagogiques. Si elle trouve que les valeurs présentées par le système sont correctes, ces dernières sont enregistrées dans son profil en tant qu'évaluation de l'objet. Dans le cas contraire, elle introduit via

l'utilitaire d'évaluation de nouvelles valeurs pour les pertinences pédagogiques des concepts annotants. Ce sera cette évaluation qui sera enregistrée dans son profil en tant qu'évaluation de l'*Objet4* et qui est donnée par :

$$\begin{pmatrix} EPr_1(Alice, Obj\acute{e}t4) \\ EPr_2(Alice, Obj\acute{e}t4) \\ EPr_3(Alice, Obj\acute{e}t4) \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

La partie dynamique de son profil est alors modifiée. Elle est donnée par : $((objet6), (objet6), (objet1, objet2, objet3, objet4))$

5.4.10 Entrepôt des objets d'apprentissage

C'est au niveau de cette structure que sont sauvegardés les objets d'apprentissage soumis par les utilisateurs. Pour la description de ces ressources nous avons opté pour un sous-schéma de LOM, représenté par les métadonnées suivantes : le titre, l'identifiant qui est un code choisi par la personne qui soumet l'objet à l'entrepôt, l'adresse URL : la combinaison de ces trois informations permet le contrôle d'unicité de l'objet soumis. Nous avons aussi inclut dans ce schéma, l'auteur de la ressource, un résumé, le domaine, le format, le type de la ressource : un tutoriel, une simulation, etc., la langue, la date de création et l'audience cible. Les composants d'ajout, de recherche et d'annotation d'objet d'apprentissage utilisent cette base pour leurs traitements.

5.4.11 Ontologie du domaine

Nous décrivons par le biais de notre ontologie, les concepts du domaine qui vont servir pour l'annotation des objets d'apprentissage (voir 6.3.1). Nous avons utilisé l'outil Protégé pour la construction de notre ontologie et pour sa traduction en format RDF/XML. Les deux composants : d'annotation et d'enrichissement de l'ontologie utilisent cette base pour leurs traitements.

5.4.12 Table des annotations

C'est la structure qui permet d'enregistrer les annotations sémantiques de l'objet d'apprentissage, introduites par l'utilisateur qui a soumis la ressource à l'entrepôt. En effet nous avons fait le choix de ne permettre l'annotation d'un objet qu'à l'utilisateur qui l'a soumis. Ce choix est dicté par notre conviction que l'utilisateur qui soumet un objet, est le plus apte à choisir dans un premier temps les concepts qui décrivent sa sémantique et d'en évaluer ensuite les degrés de couverture.

5.4.13 Table des évaluations

Cette table permet de sauvegarder les évaluations de l'utilisateur des recommandations qui lui sont présentées par le système. Elles servent ensuite dans le calcul des recommandations ultérieures pour les utilisateurs qui lui sont similaires.

5.5 Comparaison

Dans cette section, nous comparons notre recommandeur avec d'autres systèmes de recommandation d'objets d'apprentissage que nous avons décrits en 3.4. Pour ce faire, nous avons retenu les critères suivants :

- Métadonnées : Il est important de considérer pour chaque système quel schéma de métadonnées est utilisé pour la description de ses ressources.
- Gestion du démarrage à froid : Ce critère permet de savoir si le système arrive à fonctionner correctement lors de sa première utilisation.
- Algorithmes utilisés pour la recommandation : Ce critère permet de savoir comment le système procède au calcul des recommandations présentées à l'utilisateur.
- Utilisation de l'annotation sémantique : Ce critère nous informe sur l'utilisation d'autres métadonnées, pour l'amélioration des recommandations.
- Dimension de la recommandation : Généralement les systèmes de recommandations donnent des recommandations à une seule dimension, c-à-d que le système se base

sur un seul critère pour le calcul de ses recommandations. Notre système, lui propose une recommandation à plusieurs dimensions.

- Feedback de l'utilisateur : Nous voulons savoir par ce critère si l'apprenant peut interagir avec les résultats des recommandations données par le système, et de quelle façon il peut le faire et si le système prend en considération cet avis lors des évaluations suivantes.
- Feedback du système : Ce critère permet de savoir si le système fournit un feedback à l'apprenant sous n'importe quelle forme.

Tableau 5.4 – Tableau comparatif du recommandeur de LORESA avec les systèmes décrits en 3.4

<i>Systèmes</i>	<i>Métadonnées</i>	<i>Gestion démarrage à froid</i>	<i>Algorithmes</i>	<i>Dimension(s) de la recommandation</i>	<i>Utilisation de l'annotation sémantique</i>	<i>Feedback de l'utilisateur</i>	<i>Feedback du système</i>
RACOFI	Sous-ensemble de LOM	Oui, par le biais de la base de règles	Collaboratif modifié	5	Non	Évaluation des recommandations	Non
COLDEX	LOM modifié	Non	Collaboratif	1	Non	Évaluation des recommandations	Non
LORESA	sous-ensemble de LOM	Oui, par le biais de l'évaluation initiale de l'annotation sémantique	Cascade : contenu et collaboratif	3	Oui	Évaluation des recommandations	Oui, par le biais des explications des recommandations

Interprétation du tableau :

- Tous les systèmes comparés utilisent un sous-ensemble du schéma des métadonnées LOM dont la cardinalité ne dépasse pas la vingtaine d'attributs. Ceci s'explique par le fait que ce schéma qui comporte 80 descripteurs est difficile à utiliser sous sa forme intégrale.
- Comme nous l'avons vu en 3.2.3.1, les systèmes de recommandation basés sur le filtrage collaboratif souffrent du problème du démarrage à froid, à moins, de les augmenter d'une technique qui permette de gérer ce problème. Parmi les systèmes comparés notons que RACOFI et LORESA sont les deux qui permettent la gestion du démarrage à froid. Le premier système se base sur une base de règles pour générer des recommandations, même quand il y peu ou pas du tout d'évaluations. Dans notre cas nous avons utilisé les annotations sémantiques initiales avec lesquelles nous avons enrichies la description des objets d'apprentissage. Ainsi lors de la soumission de l'objet à l'entrepôt et son annotation, l'utilisateur évalue les pertinences pédagogiques de chaque concept annotants. Même si cet objet n'a pas encore été évalué, il peut être recommandé à des usagers sur la base de cette évaluation initiale.
- Algorithmes : Les systèmes RACOFI et COLDEX utilisent l'algorithme collaboratif. Dans le cas de RACOFI, cet algorithme est modifié pour en améliorer les performances (voir 3.4.1.2). Notre système utilise une technique de recommandation hybride basée sur la cascade des résultats des deux techniques de recommandation : à base de contenu et collaboratif. Nous avons opté pour cette technique par souci de convivialité, en effet même si l'utilisateur ne fait pas de demande de recommandations, le système lui présente quand même des ressources dont la sélection est basée sur les informations contenues dans le profil de l'utilisateur.
- Dimension(s) de la recommandation : Ceci nous informe sur le nombre de critères utilisés pour les évaluations des objets d'apprentissage pour la génération des recommandations. RACOFI utilise cinq critères qu'il pondère pour effectuer ses éva-

luations (voir 3.4.1), COLDEX utilise un seul critère qui est l'évaluation de l'apport du contenu de l'objet d'apprentissage au contenu global d'un cours formé par plusieurs objets. Dans LORESA, nous avons opté pour l'évaluation de trois critères qui sont les pertinences pédagogiques des trois concepts annotants au contenu de l'objet.

- Utilisation de l'annotation sémantique : Ce critère n'est utilisé que dans LORESA. En effet l'utilisation des annotations basées sur une ontologie garantit la cohérence des évaluations. Les critères utilisés dans les autres systèmes restent subjectifs et peuvent prêter à confusion tandis que dans LORESA, les critères de l'évaluation sont déterminés par l'utilisateur qui soumet l'objet en se basant sur l'ontologie chargée. Donc ce critère a la même signification pour tous les utilisateurs puisqu'il est basé sur un vocabulaire commun et formel qui peut être consulté par tous.
- Feedback de l'utilisateur : Dans les trois systèmes comparés, l'utilisateur peut donner son feedback par le biais de l'évaluation des recommandations générées.
- Feedback du système : LORESA est le seul système qui permet au système de donner un Feedback à l'utilisateur sous la forme d'explications à la demande. Ainsi si un utilisateur veut avant d'évaluer les recommandations générées par le système, en comprendre le processus de génération, il peut en faire la demande et le système lui explique comment il a procédé pour la génération de ces valeurs.

5.6 Conclusion

Au cours de ce chapitre, nous avons présenté notre architecture et expliqué la méthodologie que nous avons suivie pour la conception de notre système LORESA. Nous avons donc présenté les sous-systèmes qui constituent notre architecture avec une description détaillée de leurs modules respectifs. Le système que nous avons conçu, permet tout d'abord d'annoter sémantiquement des objets d'apprentissage par le biais d'une ontologie du domaine. Il permet ensuite et à la demande de recommander à un utilisateur, la

ressource pédagogique qui répond le mieux à ses besoins. Pour ce faire nous avons utilisé une technique qui met en cascade les deux techniques : le filtrage à base de contenu, qui prend en compte les préférences de l'utilisateur et le filtrage collaboratif qui prend en compte les usagers qui lui ressemblent. Dans le chapitre suivant, nous proposons de voir le détail de l'implémentation et la technologie utilisée dans LORESA.

CHAPITRE 6

IMPLÉMENTATION ET VALIDATION

Dans ce chapitre, nous décrivons l'environnement de développement de notre système, notamment des outils et des langages de programmation utilisés. Nous donnerons par la suite des scénarios d'utilisation de LORESA traduits par des captures d'écran. Nous abordons ensuite la validation de notre approche.

6.1 Implémentation

LORESA a été développé sous le système d'exploitation Windows XP et le serveur Web Apache Tomcat, version 2.0.54 (Fedora). Nous avons utilisé le langage de programmation PHP (PHP 5.0.4) et MySql (MySql 4.1.20) pour notre base de données. L'environnement de développement utilisé est Eclipse (3.1) avec le plugin PHP (1.1.8). L'ontologie que nous avons développée est traduite en RDF/XML pour que l'on puisse l'utiliser dans notre application PHP.

Nous donnons dans ce qui suit, une description sommaire des technologies utilisées :

- PHP : La version utilisée est PHP 5.0.4 sous le serveur Web Apache Tomcat version 2.0.54. La version 5 de PHP est dotée de *Zend Engine 2*¹. C'est ce dernier qui a véritablement introduit le modèle objet par le biais d'une gestion des erreurs basée sur le modèle des exceptions, ainsi que des fonctionnalités de gestion pour les entreprises.
- une API Document Object Model : l'API qui est disponible sous la version 5 de PHP, permet de construire une structure hiérarchique contenant des objets représentant les éléments du document, et dont les méthodes permettent d'accéder aux

¹<http://www.zend.com/engine2/ZendEngine-2.0.pdf>

propriétés. Nous avons utilisé cette API essentiellement pour le parcours et l'enrichissement de notre ontologie.

- La librairie libxslt : Celle qui est fournie sous PHP5 est libxslt version 1.1.14. Elle permet d'effectuer les transformations XSLT (transformer XML en HTML). Nous l'avons utilisée pour l'affichage de l'ontologie dans notre application Web.
- MySql : La version qui est utilisée conjointement avec PHP 5 est la version 4.1.20.

Notre application est basée sur l'architecture client-serveur où la communication est réalisée à travers un dialogue entre un client et un serveur. Le mode d'interaction est basée sur la paire (requête, réponse) où le client est l'instigateur de la requête. Le serveur qui reçoit cette requête, la traite et fournit une réponse au client. Ainsi une application basée Web doit permettre ces échanges d'informations entre le client représenté par le navigateur de l'utilisateur, et le serveur représenté par un serveur Web. Plusieurs protocoles peuvent être utilisés pour permettre ces échanges parmi lesquels le protocole HTTP (HyperText Transfer Protocol). Lorsque le client saisit une URL (Uniform Resource Locator) dans son navigateur, une requête http est envoyée au serveur Web, c'est-à-dire que le client interroge le serveur Web. Le serveur Web renvoie ensuite une réponse basée également sur le protocole http.

6.2 Utilisation de LORESA

Comme nous l'avons vu décrit en 5.2.1, LORESA est le couplage d'un *système de recommandation* à un *système d'annotation* appliqué à un *entrepôt d'objets d'apprentissage*. Pour illustrer ses fonctionnalités de nous présentons divers scénarios d'utilisation sous forme de captures d'écrans. Il est important de noter que pour un utilisateur de LORESA, le passage d'un sous-système à un autre se fait de manière totalement transparente : il peut vouloir ajouter² une ressource à l'entrepôt ensuite rechercher des objets qui répondent à certains critères et demander des recommandations³. Il peut ensuite passer à

²Entrepôt d'objets d'apprentissage

³Système de recommandation

l'annotation⁴ de l'objet ajouté en consultant d'abord l'ontologie de domaine.

La porte d'entrée de notre système est la page d'accueil de LORESA (voir figure 6.1) qui donne une brève description du système, de ses objectifs ainsi que de ses principales fonctionnalités. C'est à ce niveau qu'un utilisateur peut s'enregistrer en cliquant sur

LORESA La recommandation sémantique des objets d'apprentissage

LORESA (Learning Objects REcommender system based on Semantic Annotation)

Ce système **recommande** à un usager enregistré des objets d'apprentissage décrits par leurs métadonnées et leurs **annotations sémantiques**.

L'annotation sémantique est basée sur l'utilisation d'un **vocabulaire contrôlé** et sert à définir l'objectif pédagogique d'un objet d'apprentissage en terme de **concepts couverts**. Dans LORESA, chaque objet est décrit par trois concepts : un principal et deux secondaires. L'annotateur de l'objet attribue à chaque concept une valeur entre 1 et 5 qui décrit la pertinence pédagogique de l'objet pour l'acquisition de chaque concept, où 1 : signifie une faible pertinence pédagogique et 5 signifie une forte pertinence pédagogique.

Les étapes d'utilisation de notre système suivent le schéma suivant :

- Un utilisateur commence par s'enregistrer.
- Dès qu'il s'identifie, LORESA lui présente des ressources (objets) qui correspondent à son profil initial, quand celles-ci existent dans l'entrepôt.
- L'utilisateur a le choix d'approfondir sa recherche en choisissant de faire une recherche sémantique ou faire une autre recherche.
- LORESA, présente ensuite ses recommandations à l'utilisateur en se basant sur les évaluations des autres usagers du système.
- L'utilisateur est ensuite appelé à évaluer ces recommandations pour améliorer la précision de LORESA. A la fin , l'utilisateur est sollicité à évaluer globalement LORESA en cliquant sur le bouton **EVALUER**

Si c'est la première fois que vous utilisez notre système, veuillez vous **enregistrer**, sinon veuillez vous identifier. Merci

Figure 6.1 – Page d'accueil de LORESA

le lien *Enregistrez-vous*. L'activation de ce lien conduit à l'affichage de la page *enregistrement d'un utilisateur* (voir figure 6.2) où sont décrites les informations requises à cette opération parmi lesquelles : pseudonyme, mot de passe, champs d'intérêt, etc.

En cliquant sur le bouton *Continuer*, le formulaire d'inscription s'affiche et l'utilisateur peut introduire les informations demandées qui sont contrôlées à divers niveaux (voir figure 6.3).

À l'issue de cette opération, l'utilisateur peut s'authentifier auprès du système en appuyant sur le bouton *Connecter*, ce qui lui garantit alors l'accès aux fonctionnalités du système.

⁴Système d'annotation

LORESA La recommandation sémantique des objets d'apprentissage

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation

Enregistrement d'un nouvel utilisateur

Bonjour, cette opération vous permet d'enregistrer **votre profil** dans notre système. Ces informations sont utilisées dans un premier temps pour vous **recommander** des ressources pédagogiques

Pour se faire vous êtes invités à remplir les champs suivants:

Un pseudonyme,
que vous utiliserez dans notre système;

Un mot de passe,
d'une longueur minimale de 04 caractères;

Un profil,
Enseignant ou apprenant;

Un champs d'intérêt,
parmis ceux proposés;

Votre format,
de ressources préférées;

Votre Langue,
de ressources préférées;

Une adresse courriel valide,

Vous avez un compte ?
Identifiez-vous

Nom d'utilisateur

Mot de passe

Pas encore de compte ?
Enregistrez-vous!!

Figure 6.2 – Page d'explication de la procédure d'enregistrement

LORESA La recommandation sémantique des objets d'apprentissage

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation

Nouvel utilisateur

Nom :

mot de passe :

Confirmez votre mot de passe :

Rôle : Apprenant

Champ d'intérêt :

Le format préféré :

La langue préférée : Anglais
 Français
 Espagnol
 Toutes

Courriel :

Vous avez un compte ?
Identifiez-vous

Nom d'utilisateur

Mot de passe

Pas encore de compte ?
Enregistrez-vous!!

Figure 6.3 – Enregistrement de l'utilisateur *Jean*

6.2.1 LORESA : système de recommandation

Nous allons présenter dans cette section le scénario d'un utilisateur qui veut obtenir des recommandations de ressources pédagogiques.

6.2.1.1 La recommandation

Dès que l'utilisateur s'authentifie, le *recommandeur* de LORESA applique un filtrage à base de contenu et lui présente des ressources pédagogiques qui correspondent à ses préférences, des informations contenues dans son profil (voir section 5.4.9). Dans ce cas l'utilisateur a le choix de faire une recherche sémantique des ressources affichées en cliquant sur le lien *recherche sémantique* ou de procéder à une nouvelle recherche en cliquant sur le lien *autre recherche* (voir figure 6.4).

Objets d'apprentissage | Profil utilisateur | Ontologies | Évaluation

Bienvenu (e), jean

Utilisateur : jean
Rôle : Apprenant

Selon les informations introduites lors de votre enregistrement, le système vous recommande les ressources suivantes :

Identifiant	Titre	Format	Lien
NETH02	Carrier Sense Multiple Access/Collision Detection (CSMA/CD)	PDF	http://www.org.abdn.ac.uk/users/gorry/course/lan-pages/csma-cd.html
NTCP1P	Introduction to the TCP/IP LAN	PDF	http://handsonbowto.com/lan101.html

Vous pouvez approfondir votre recherche en cliquant sur *recherche sémantique*. Cette option vous permet d'afficher les annotations sémantiques des objets recherchés.

Dans LORESA : Ceci consiste à afficher les concepts traités par l'objet, un **concept principal** et deux **concepts secondaires**.

Vous pouvez également faire une *autre recherche*.

Figure 6.4 – Ressources recommandées à Jean lors de son authentification

Si l'utilisateur clique sur le bouton *recherche sémantique*, Les annotations sémantiques des objets présentés sont affichées. Comme nous l'avons vu, ces annotations sont en fait des concepts du domaine, plus précisément un concept principal et deux concepts secondaires que cette ressource couvre à des degrés donnés.

En prenant connaissance des concepts couverts par les ressources présentées, deux choix s'offrent à l'utilisateur il peut obtenir à la demande des recommandations concernant les objets qu'il sélectionne ou faire une nouvelle recherche (voir figure 6.5).

En cliquant sur le bouton *recommandation(s)*, le système présente à Jean les recommandations des pertinences pédagogiques présentées sous forme de valeurs numériques entre 1 et 5 (voir figure 6.6).

LORESA La recommandation sémantique des objets d'apprentissage

Objets d'apprentissage | Profil utilisateur | Ontologies | Évaluation

Rechercher | Ajouter | Annoter | Retour

Usager : jean
Rôle : Apprenant

Recherche sémantique d'o

Pour vous aider à choisir les objets qui vous conviennent le mieux, vous êtes invités à cliquer sur le bouton **Recommandations**.

Vous pouvez également faire une autre recherche

Identifiant	Concept Principal	Concept Secondaire 1	Concept Secondaire 2
NETH02	Circuit-switching networks	Frame relay networks	Packet-switching networks <input type="checkbox"/>
NTCPIP	Applications (SMTP, FTP, etc.)	Routing protocols	Protocol verification <input checked="" type="checkbox"/>

Recommandation(s) Autre Recherche Retour

Figure 6.5 – Les annotations sémantiques des objets présentées à Jean

Objets d'apprentissage | Profil utilisateur | Ontologies | Évaluation

Rechercher | Ajouter | Annoter | Retour

Usager : jean
Rôle : Apprenant

Après avoir sélectionné les objets dont les concepts vous intéressent, LORESA vous présente **ses recommandations** sous la forme de : valeurs de pertinence pédagogique de chaque concept couvert sur une échelle de 1 à 5, où 1 signifie une **faible pertinence pédagogique** et 5 une **forte pertinence pédagogique**. Ces valeurs ont été évaluées par les usagers de notre système.

LORESA ne prend en compte que les évaluations des usagers qui sont proches de vous. Si vous voulez avoir des **explications de toutes les recommandations données**, vous pouvez cliquer sur le bouton **Explication**.

Pour améliorer la qualité de nos recommandations, nous vous demandons de **cocher** les objets, dont les recommandations ne vous conviennent pas cliquer sur le bouton **Evaluation**.

Identifiant	Concept Principal	PPG	Concept Secondaire 1	PPG	Concept Secondaire 2	PPG	Lien de la ressource
NTCPIP	Applications (SMTP, FTP, etc.)	5.0	Routing protocols	2.0	Protocol verification	2.0	http://handsonhowto.com/lan101.html <input checked="" type="checkbox"/>

PPG : Pertinence Pédagogique, E : Objet déjà évalué.

Evaluation Retour Explication

Figure 6.6 – La recommandation de l'objet sélectionné

En prenant connaissance de ces recommandations et en consultant l'objet en question (voir figure 6.7), l'utilisateur peut s'il le désire obtenir des explications concernant les recommandations calculées en cliquant sur le bouton *explication* (voir section 5.3.4.3).

Cette possibilité constitue une de nos contributions dans ce travail, le processus de

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation

Rechercher | Ajouter | Annoter | Retour

Utilisateur : Jean
Rôle : Apprenant

Identifiant	Concept Principal	PPG	Concept Secondaire 1	PPG	Concept Secondaire 2	PPG	Lien de la ressource
NTCPIP	Applications (SMTP, FTP, etc.)	5.0	Routing protocols	2.0	Protocol verification	2.0	http://handsonhowto.com/lan101.html <input checked="" type="checkbox"/>

PPG : Pertinence Pédagogique, E : Objet déjà évalué.

Evaluation | Retour | Explication

Hardware
Computer Systems
Software
Data
Information Systems

The TCP/IP Network : Hands-on How-to - Windows Internet Explorer

A LAN: <http://handsonhowto.com/2007/lan101/>
this room, give or take the dozen down the hall * it implies a fast connection — dozens, hundreds, even thousands of times faster than a dial-up modem. It used to be pretty challenging to hook computers together this way, but the parts got cheaper and the software got smarter

TCP/IP is one of the reasons it's easier. TCP/IP is actually two things: TCP, or Transmission Control Protocol, and IP, or Internet Protocol. (A protocol is just "a clearly defined way of doing something.") TCP/IP (the combination of TCP with IP) defines one way that computers on a network can communicate by exchanging packets. One definition of "The Internet" with a capital "I" is simply "all the computers that connect to each other and use this...

Figure 6.7 – Consultation de l'objet

recommandation devient alors un processus totalement transparent à l'utilisateur puisque l'explication permet de donner à l'utilisateur les informations qui ont servi au calcul des recommandations qui lui ont été présentées (voir figure 6.8).

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation

Rechercher | Ajouter | Annoter | Retour

NTCPIP	Recom/calculée	Utilisateur similaire/Annotateur	Evaluation
Non	sihem	Princ Applications (SMTP, FTP, etc.) : 5, Secon1 Routing protocols : 2, Secon2 Protocol verification : 2	

Princ, Secon1 et Secon2 : Concepts principal, secondaire 1 et secondaire 2 respectivement.
Recom calculée (Oui/Non) : Oui, les valeurs recommandées sont calculées par LORESA.
Non, les valeurs recommandées sont les valeurs des annotations initiales données par l'annotateur de l'objet.

Retour | Autre recherche

Figure 6.8 – Explication de la recommandation

Dans ce cas précis, même si Jean qui est un nouvel utilisateur, n'ayant donc pas encore évalué de recommandations, obtient quand même des recommandations. Loresa remédie donc au problème du *démarrage à froid*. Dans l'explication, il y a le champs *recommen-*

dation calculée qui peut prendre les valeurs *oui* ou *non*. Ce champs renseigne l'utilisateur sur le statut des valeurs des recommandations affichées. Si ces valeurs sont calculées (*oui*), ceci indique que le recommandeur a maintenant appliqué le filtrage collaboratif (c-à-d que ces conditions sont vérifiées : des voisins de Jean ont déjà évalué cet objet). Dans le cas contraire, ceci indique que ces valeurs ne sont pas calculées (*non*) et sont données initialement par l'annotateur de l'objet. Le prochain champ indique selon la valeur du champ précédent (*oui* ou *non*), l'usager qui a contribué à donner cette recommandation : un usager similaire ou l'annotateur de l'objet.

6.2.1.2 Évaluation de la recommandation

Après avoir consulté l'objet, si l'utilisateur n'est pas d'accord avec les valeurs des pertinences pédagogiques données par le système, il peut évaluer les recommandations présentées par le système en cliquant sur le bouton *Evaluation*. Une page s'affiche à l'utilisateur qui lui explique les objectifs de la procédure d'évaluation, ainsi que la liste des objets qu'il veut évaluer (voir figure 6.9).

Figure 6.9 – Choix de l'objet à évaluer

Lorsque l'usager sélectionne les objets qu'il veut évaluer, le système lui affiche les valeurs des pertinences pédagogiques qui lui ont été recommandées et lui demande, lesquelles il veut modifier (voir figure 6.10).

Évaluation des annotations sémantiques

LORESA vous a recommandé pour l'objet NTCPIP ce qui suit :

1. Concept principal : **Applications (SMTP, FTP, etc.)** une pertinence de : **5.0**
 Trouvez-vous la valeur de la pertinence pédagogique de cette annotation adéquate ? Oui
 Non

2. Concept secondaire 1: **Routing protocols** d'une pertinence de : **2.0**
 Trouvez-vous la valeur de la pertinence sémantique de cette annotation adéquate ? Oui
 Non

3. Concept Secondaire 2 : **Protocol verification** d'une pertinence de : **2.0**
 Trouvez-vous la valeur de la pertinence sémantique de cette annotation adéquate ? Oui
 Non

Figure 6.10 – Choix des concepts à évaluer

En sélectionnant les concepts dont il veut modifier la pertinence, le système lui affiche un formulaire où il peut introduire les nouvelles valeurs correspondantes aux concepts qu'il a sélectionnés et qui reflètent à son avis mieux leurs pertinence pédagogiques (voir figure 6.11).

Objets d'apprentissage | Profil utilisateur | Ontologies | Évaluation

Rechercher | Ajouter | Annoter | Retour

Évaluation des annotations sémantiques

Objet : NTCPIP

Vous n'êtes pas en accord avec la valeur de la pertinence pédagogique du concept principal : **Applications (SMTP, FTP, etc.)**
 Quelle valeur souhaitez-vous accorder à la pertinence du concept principal ?

Vous n'êtes pas en accord avec la valeur de la pertinence pédagogique du concept secondaire 1 : **Routing protocols**
 Quelle valeur souhaitez-vous accorder à la pertinence du concept secondaire 1 ?

Vous n'êtes pas en accord avec la valeur de la pertinence pédagogique du concept secondaire 2 : **Protocol verification**
 Quelle valeur souhaitez-vous accorder à la pertinence du concept secondaire 2 ?

Figure 6.11 – Les nouvelles valeurs des pertinences pédagogiques

En introduisant ces nouvelles valeurs, le profil de l'utilisateur est automatiquement modifié pour prendre en compte ces informations. L'utilisateur peut ainsi consulter son profil et en modifier certaines informations, ceci est expliqué en détails dans la section suivante.

6.2.1.3 Consultation et modification du profil

L'utilisateur peut à tout moment consulter son profil et ce en appuyant d'abord sur le bouton *Profil Utilisateur* et ensuite sur le bouton *Consulter*. A cet effet, il obtient une page qui l'informe sur les opérations qu'il a effectuées dans le système depuis son enregistrement (voir figure 6.12).

The screenshot shows a web interface for user profile management. At the top, there are navigation tabs: 'Objets d'apprentissage', 'Profil utilisateur', 'Ontologies', and 'Evaluation'. Below these are buttons for 'Consulter', 'Modifier', and 'Retour'. The main content area displays a greeting 'Bonjour, jean' and a message: 'cette procédure vous permet de consulter les caractéristiques dynamiques de votre profil : vous avez évalué et les usagers dont le profil est proche du votre.' Below this, there are statistics: 'Vous êtes enregistré dans Loresa depuis : 2007-11-13' and 'Depuis Votre enregistrement :'. A list of statistics follows: 'vous avez soumis : 0 objet(s) d'apprentissage', 'vous avez annoté : 0 objet(s) d'apprentissage', and 'vous avez évalué : 1 objet(s) d'apprentissage.' At the bottom, there is a section for 'Usager(s) similaire (s) :' and a 'Retour' button. On the left side, there is a sidebar with a 'Nouveautés' section showing '100 Objets' and '68 Usagers', and a 'Consulter' button. Below that, there are categories: 'Hardware', 'Computer Systems', and 'Software'.

Figure 6.12 – Profil de Jean

Nous voyons bien que Jean, depuis son enregistrement n'a pas soumis d'objets d'apprentissage, il n'en a pas annoté mais il a évalué un seul objet en l'occurrence l'objet *NTCPIP*. Notons que pour le moment, vu que Jean n'a évalué un objet qui n'a pas été évalué par d'autres utilisateurs, son voisinage est vide.

Au fur et à mesure que Jean évalue les recommandations que lui présente LORESA, des usagers intègrent son voisinage et vice versa. Après avoir effectué certaines évaluations nous obtenons le nouveau profil de Jean calculé par le système (voir figure 6.13).

Cependant l'utilisateur a la possibilité de modifier manuellement certaines données de son profil, plus précisément les données statiques : mot de passe, rôle, domaine d'intérêt (voir figure 6.14).

LORESA La recommandation sémantique des objets d'apprentissage

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation

Consulter | Modifier | Retour

Usager : Jean
Rôle : Apprenant

Bonjour, Jean
cette procédure vous permet de consulter les caractéristiques **dynamiques** de votre profil :
vous avez évalué et les usagers dont le profil est **proche** du votre.

Vous êtes enregistré dans Loresa depuis : 2007-11-13

Depuis votre enregistrement :

- vous avez soumis : 0 objet(s) d'apprentissage.
- vous avez annoté : 0 objet(s) d'apprentissage.
- vous avez évalué : 3 objet(s) d'apprentissage.

Usager(s) similaire(s) :

- zakia

Retour

Figure 6.13 – Nouveau profil de Jean

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation

Consulter | Modifier | Retour

Usager : Jean
Rôle : Apprenant

Bonjour, Jean
cette procédure vous permet de modifier les éléments **statiques** de votre profil :

- Votre mot de passe ;
- Votre profil ;
- Votre champs d'intérêt ;
- Votre format préféré.

Votre nouveau mot de passe : ****

Confirmez votre nouveau mot de passe : ****

Votre Profil : Apprenant Votre nouveau profil : Enseignant

Votre champs d'intérêt : Computer Systems Organization Votre nouveau champs d'intérêt : Hardware

Votre format préféré : PDF Votre nouveau format : Autre

Envoyer Retour

Figure 6.14 – Modification du profil de Jean

6.2.2 LORESA : l'entrepôt d'objets d'apprentissage et le système d'annotation

6.2.2.1 Ajout d'un objet d'apprentissage à l'entrepôt

Dans ce mode un usager enregistré peut vouloir soumettre des objets d'apprentissage à l'entrepôt. Pour ce faire il doit cliquer sur le bouton *Objets d'apprentissage* et *Ajouter*. Ceci permet d'ajouter un objet d'apprentissage au dépôt en deux étapes. La première consiste à inviter l'usager à introduire : le titre de l'objet, son URL et un identifiant qui est un chaîne alphanumérique pour désigner l'objet. Les deux premiers champs sont utilisés

comme clé primaire afin de garantir l'unicité de l'objet dans l'entrepôt d'apprentissage (voir figure 6.15)

LORESA La recommandation sémantique des objets d'apprentissage

Objets d'apprentissage | Profil utilisateur | Ontologies | Évaluation

Réchercher | Ajouter | Annoter | Retour

Usager : jean
Rôle : Apprenant

Ajout d'un objet d'apprentissage

Cette procédure vous permet de soumettre un objet d'apprentissage à l'entrepôt
(*): Champs obligatoires.

Titre (*):

URL (*):

Identifiant (*): (Une chaîne alphanumérique de 06 caractères)

Figure 6.15 – Première étape d'enregistrement d'un objet d'apprentissage

Après vérification de l'unicité de l'objet d'apprentissage, l'utilisateur est invité à remplir les autres champs de description de cette ressource. Comme nous l'avons décrit, nous

Ajout d'un objet d'apprentissage

Auteur :

Domaine :

Résumé du contenu :

Type de la ressource :

Format de la ressource :

Audience cible : Secondaire Universitaire (1 er cycle) Universitaire (2 et 3 cycles) Professionnel

Date de création : (AAAA-MM-JJ)

Langue : Anglais Français Espagnol Autre

Difficulté : 1 2 3 4

Figure 6.16 – Étape finale de l'enregistrement d'un objet d'apprentissage

avons choisi d'utiliser un sous-ensemble du schéma LOM, que nous avons décrits par les champs : Auteur, domaine, résumé du contenu, type de la ressource, format, audience cible, date de création, langue et difficulté (voir figure 6.16).

6.2.2.2 Annotation de l'objet d'apprentissage

À l'issue de la phase d'enregistrement, LORESA demande à l'utilisateur s'il veut *annoter* son objet d'apprentissage (voir figure 6.17). Si l'usager choisit d'annoter son objet

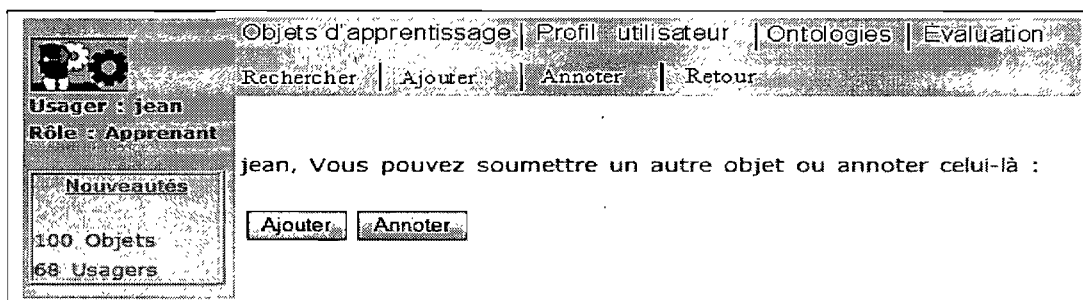


Figure 6.17 – Demande d'annotation de l'objet

tout de suite (un choix recommandé par le système), LORESA affiche à l'utilisateur une page descriptive qui permet de lui expliquer cette procédure (voir figure 6.18). Cependant,

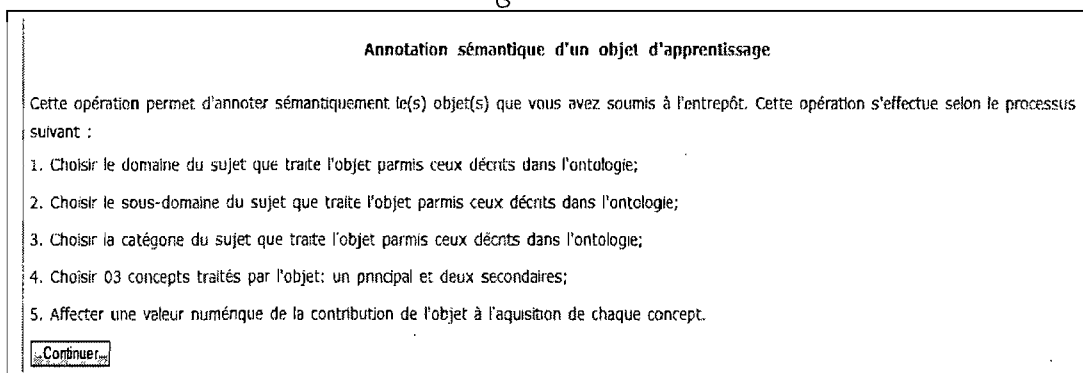


Figure 6.18 – Explication de la procédure d'annotation

l'utilisateur peut choisir de ne pas annoter son objet immédiatement après l'avoir soumis. Il peut quand même le faire de façon différée en cliquant sur le bouton *Objets d'apprentissage* ensuite sur le bouton *Annoter*. C'est en annotant l'objet, que ce dernier peut être recommandé à d'autres usagers. Si l'objet n'est pas annoté il ne peut être trouvé que par le biais d'une recherche classique. En cliquant sur le bouton *Continuer*, l'utilisateur est invité à choisir le sous-domaine du sujet de son objet (voir figure 6.19).

Annotation sémantique d'un objet d'apprentissage

Identifiant : DNSC01

Domaine : Computer Systems Organization

Choisissez un sous domaine :

GENERAL

PROCESSOR ARCHITECTURES

COMPUTER-COMMUNICATION NETWORKS

SPECIAL-PURPOSE AND APPLICATION-BASED SYSTEMS

PERFORMANCE OF SYSTEMS

COMPUTER SYSTEM IMPLEMENTATION

MISCELLANEOUS

Figure 6.19 – Choix du champs sous-domaines de l'annotation

Ces informations sont extraites à partir de l'ontologie, qui, d'après le domaine introduit lors de la soumission de l'objet à l'entrepôt, permettent d'afficher les possibilités pour le sous-domaine et la catégorie (voir figure 6.20).

Annotation sémantique d'un objet d'apprentissage

Identifiant : DNSC01

Domaine : Computer Systems Organization

Sous domaine : COMPUTER-COMMUNICATION

Catégorie :

COMPUTER-COMMUNICATION NETWORKS

General

Network Architecture and Design

Network Protocols

Network Operations

Distributed Systems

Figure 6.20 – Choix de la catégorie de l'annotation

LORESA affiche maintenant les concepts définis dans l'ontologie sous les sous-domaine et catégorie choisis. L'utilisateur choisit alors un concept principal et deux concept secondaires avec des valeurs de pertinences pédagogiques pour chaque concept (voir figure 6.21).

Ceci achève l'annotation sémantique de cet objet qui peut ainsi être recommandé lors de recherches sémantiques réalisées par d'autres usagers (voir figure 6.22).

Annotation sémantique d'un objet d'apprentissage

Identifiant : DNSCO1
 Domaine : Computer Systems Organization
 Sous domaine : COMPUTER-COMMUNICATION
 Catégorie : Network Protocols

Concepts traités par l'objet :

Concept principal : Valeur pertinence pédagogique :

Concept Secondaire 1 : Valeur pertinence pédagogique :

Concept secondaire 2 : Valeur pertinence pédagogique :

Figure 6.21 – Choix des concepts et des pertinences pédagogiques

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation

Rechercher | Ajouter | Annoter | Retour

Usager : lina
Rôle : Apprenant

Recherche sémantique d'objets

Pour vous aider à choisir les objets qui vous conviennent le mieux, vous êtes invités à choisir ceux qui couvrent les critères sur le bouton **Recommandations**.

Vous pouvez également faire une autre recherche

Identifiant	Concept Principal	Concept Secondaire 1	Concept Secondaire 2
NETW04	Buses	Ethernet (e.g., CSMA/CD)	Internet (e.g., TCP/IP)
NETH01	Centralized networks	Network communications	Store and forward networks
DNSCO1	IP addressing	Applications (SMTP, FTP, etc.)	Protocol architecture (OSI model)

Figure 6.22 – Recommandation possible de l'objet annoté

6.2.2.3 Recherche de ressources pédagogiques

Cette fonctionnalité peut être obtenue de diverses façons parmi lesquelles, en appuyant sur le bouton *Objets d'apprentissage* ensuite sur le bouton *Rechercher* (voir figure 6.23).

Ceci permet d'afficher le formulaire où l'utilisateur est invité à introduire les valeurs des critères de sa recherche. Le système affiche à l'utilisateur les ressources qui répondent aux critères spécifiés dans son formulaire de recherche (voir figure 6.24)

Figure 6.23 – Recherche de ressources pédagogiques

Identifiant	Titre	Format	Lien	Date de création
VPNETW	Virtual Private Networks	HTML	http://www.iec.org/online/tutorials/vpn/	2002-01-31
PROT02	TCP/IP (Transmission Control Protocol/Internet Protocol)	HTML	http://www.cerigside.com/tcpip.asp	2002-02-02
PROT10	IP Addressing and Subnetting	Texte	http://www.learnosubnet.com/	2002-10-02
TCP15	IP Addressing and Subnetting for New Users	Texte	http://www.cisco.com/warp/public/701/3.html	2002-10-02
CLISER	Client/Server Software Architectures - An Overview	HTML	http://www.sei.cmu.edu/str/descriptions/clientserver_body.html	2002-10-24
NTCPIP	Introduction to the TCP/IP LAN	PDF	http://haedsonbowlo.com/lan101.html	2003-04-06

Figure 6.24 – Affichage des ressources pédagogiques répondants aux critères

Il peut faire une autre recherche, en spécifiant d'autres critères et ce en appuyant sur le bouton *Autre recherche*. Le deuxième choix qui s'offre à lui et de continuer avec les résultats de cette recherche et ce en appuyant sur le bouton *Recherche sémantique* (voir 6.2.1.3)

6.2.2.4 Consultation des ressources

Cette fonctionnalité permet de consulter le contenu de l'entrepôt en sélectionnant le domaine recherché dans la barre se trouvant à la gauche de la page d'accueil (voir figure 6.25).

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation

Rechercher | Ajouter | Annoter | Retour

Usager : jean
Rôle : Apprenant

Nouveautés

100 Objets
69 Usagers

Consulter as

Hardware

Computer Systems

Software

Data

Information Systems

Déconnexion

Asynchronous Transfer Mode (ATM) Fundamentals Tutorial
Auteur : Anna Gibson
Résumé : ATM is a technology that will enable carriers to capitalize on a number of revenue opportunities
Format : HTML
date création : 2001-03-11

Introduction to the TCP/IP LAN
Auteur : Therezita Ortiz
Résumé : Very simple explanation on how to set up a TCP/IP network (LAN).
Format : PDF
date création : 2003-04-06

Virtual Private Networks
Auteur : Judy Ann Serwatka
Résumé : Tutorial and quiz on virtual private networks.
Format : HTML
date création : 2002-01-31

Client/Server Software Architectures--An Overview
Auteur : Darleen Sadoski
Résumé : Client/Server software architectures concepts description.
Format : HTML
date création : 2002-10-24

Resultats : page 2 de 3

Précédent Suivant

Figure 6.25 – Consultation des ressources du domaine *Computer Systems*

6.2.3 LORESA : l'ontologie du domaine

Pour la construction de l'ontologie de LORESA, nous avons utilisé protégé. Nous l'avons ensuite traduite en RDF/XML via un utilitaire fourni par cette plateforme. LORESA offre deux modes d'interaction avec l'ontologie : le mode consultation pour tous les usagers et le mode ajout pour un seul usager ayant le profil : Expert.

6.2.3.1 Consultation de l'ontologie de LORESA

Ce mode permet à un utilisateur enregistré de consulter l'ontologie utilisée en appuyant sur le bouton : *Ontologies* ensuite sur le bouton *Consulter* (voir figure 6.26). Cette

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation

Consulter | Ajouter | Retour

Usager : jean
Rôle : Apprenant

Nouveautés
100 Objets
68 Usagers

Consulter »

Hardware
Computer Systems
Software

Consultation de l'ontologie utilisée
Cette procédure vous permet de consulter l'ontologie utilisée et ce en vous permettant de parcourir les domaines généraux de cette ontologie. Si vous voulez consulter les sous domaines, cliquer sur le vous intéresse, sous domaines, catégories et concepts.

Code domaine	Intitulé	Consulter	
B	Hardware	<input type="radio"/>	
C	Computer Systems Organization	<input checked="" type="radio"/>	
D			
	Code sous-domaine	Intitulé	
E	C.0	GENERAL	<input type="radio"/>
H	C.1	PROCESSOR ARCHITECTURES	<input type="radio"/>
	C.2	COMPUTER-COMMUNICATION NETWORKS	<input checked="" type="radio"/>

Figure 6.26 – Consultation de l'ontologie

option permet d'afficher pour chaque choix de l'utilisateur du domaine, du sous-domaine et de la catégorie, les concepts enregistrés dans cette hiérarchie. Il peut alors consulter les concepts (voir figure 6.27).

6.2.3.2 Ajout de concepts à l'ontologie de LORESA

Cette option permet à un seul utilisateur en l'occurrence l'utilisateur ayant le profil : *Expert* d'enrichir l'ontologie en concepts (voir figure 6.28).

En effet comme nous l'avons expliqué en 5.3.1, notre ontologie est basée sur l'enrichissement de la classification de l'ACM à un degré plus élevé de granularité en ajoutant des concepts sous l'arborescence adéquate. Ceci requiert évidemment un niveau d'expertise qui garantit la pertinence de l'ajout c.-à-d. que ce qui a été ajouté peut vraiment être considéré comme un concept mais aussi la cohérence de son emplacement. À l'issue de cette

LORESA La recommandation sémantique des objets d'apprentissage

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation
 Consulter | Ajouter | Retour

Usager : Jean
 Rôle : Apprenant

Consultation des catégories

Code	Catégorie	Intitulé	Consulter
C.2.0		General	<input type="radio"/>
C.2.1		Network Architecture and Design	<input type="radio"/>
C.2.2		Network Protocols	<input checked="" type="radio"/>
C.2.3		Network Operations	<input type="radio"/>
C.2.4		Distributed Systems	<input type="radio"/>
C.2.5		Local and Wide-Area	<input type="radio"/>
C.2.6		Internetworking	<input type="radio"/>
C.2.m		Miscellaneous	<input type="radio"/>

Intitulé Concept

Applications (SMTP, FTP, etc.)
Protocol architecture (OSI model)
Protocol verification
Routing protocols
point-to-point tunneling protocol (PPTP)
layer-2 forwarding (L2F)

Consulter... Retour...

Figure 6.27 – Consultation des concepts

LORESA La recommandation sémantique des objets d'apprentissage

Objets d'apprentissage | Profil utilisateur | Ontologies | Evaluation
 Consulter | Ajouter | Retour

Usager : sihem
 Rôle : Expert

Enrichissement de l'ontologie

Le domaine à enrichir : Hardware
 Computer Systems Organization
 Software

Domaine à enrichir : C

Choisissez le sous domaine : COMPUTER-COMMUNICATION NETWORKS

Catégorie : Network Protocols

Introduisez le Concept que vous voulez ajouter : Simple Network Managment Protocol

Envoyer

Figure 6.28 – Ajout d'un concept à l'ontologie de LORESA

opération, le concept ajouté par l'utilisateur *Expert* peut être consulté et utilisé lors des annotations de nouveaux objets.

Ceci achève la partie implémentation où nous avons passé en revue toutes fonctionnalités de LORESA, nous passons ensuite à la validation de notre système sujet de notre prochaine section.

6.3 Validation

Afin de valider notre système, nous avons conduit une étude empirique par laquelle nous voulions vérifier les hypothèses suivantes :

1. Étant donné que LORESA est destiné aussi bien aux enseignants qui veulent chercher une ressource pour construire un cours, qu'aux apprenants qui veulent renforcer leurs connaissances en consultant des ressources supplémentaires. Nous avons voulu mesurer l'utilisabilité de notre système pour ces deux populations en émettant l'hypothèse suivante : *les enseignants sont plus enclins à utiliser les fonctionnalités d'ajout et d'annotation de l'entrepôt d'objets d'apprentissage, que ne le sont les apprenants, alors que ces derniers sont plus intéressés par la recherche de ressources.* Concernant les autres fonctionnalités nous n'avions pas d'à-prioris et nous voulons seulement nous renseigner sur leurs pertinences et jusqu'à quel point sont-elles utilisées. Pour des besoins de clareté, nous utilisons le mot : profil pour désigner le rôle de l'utilisateur à ne pas confondre avec le profil utilisateur de LORESA.
2. Concernant le choix de la technique de la recommandation utilisée nous avons émis l'hypothèse suivante : *l'application de la technique hybride du filtrage donne de meilleurs résultats que l'application de chaque technique toute seule.*

Nous avons ensuite analysé les données recueillies du formulaire d'évaluation globale du système que les participants sont invités à remplir avant de se déconnecter du système. Cette évaluation nous informe entre autres sur l'appréciation par les utilisateurs de l'apport de l'annotation sémantique, l'apport de l'explication de la recommandation, etc.

6.3.1 Contexte de la validation

Nous avons conduit notre expérimentation sur une population de 68 étudiants des cycles supérieurs de l'institut d'informatique et de recherche opérationnelle de l'université de Montréal pendant 10 jours. Nous avons fait ce choix pour nous assurer d'obtenir les deux profils d'utilisateurs de notre système : enseignants et apprenants étant donné que

certaines étudiants de la maîtrise ou du doctorat assurent la fonction d'enseignants auprès d'étudiants du premier cycle. Leur répartition selon le profil est donnée par le tableau 6.1.

Tableau 6.1 – Tableau Descriptif du contexte de la validation de LORESA

Nombre de participants	68
Profils des participants	19 Enseignants et 49 apprenants
Feedback des utilisateurs	explicite
Nombre des champs à évaluer	3 au maximum
Type d'évaluation	valeur entre 1 et 5
Domaine des objets évalués	Ressources Web en informatique
Nombre des ressources évaluées	100 ressources couvrant divers domaines

Pour le deuxième volet de la validation, nous avons opté pour le partage de cette même population aléatoirement en trois catégories : la première recevait des recommandations basées sur le filtrage par contenu (CT dans le tableau, nous avons désactivé la technique du filtrage collaboratif), la deuxième catégorie avait accès à des recommandations basées uniquement sur le filtrage collaboratif (CO dans le tableau, nous avons désactivé à cet effet, la technique du filtrage par contenu) et la dernière avait des recommandations basées sur une combinaison des deux techniques (HY dans le tableau).

- Les participants : 68 étudiants des études supérieures de l'institut d'informatique et de recherche opérationnelle de l'université de Montréal pendant une période de 10 jours environ.
- Le profil : Les étudiants ayant participé à notre expérimentation sont classés en deux catégories : apprenants et enseignants (l'information : rôle recueillie lors de leurs enregistrement). La répartition des participants est donnée par le tableau 6.2. Nous utilisons l'information du profil pour la validation des deux hypothèses que nous avons décrits en 6.3. Nous donnons en figure 6.29, le graphe de la répartition des participants selon le rôle et la technique de recommandation utilisée.

Tableau 6.2 – Répartition selon la technique de recommandation utilisée et selon le profil

	CT	CO	HY	Total
Enseignants	5	6	8	19
Apprenants	19	18	12	49

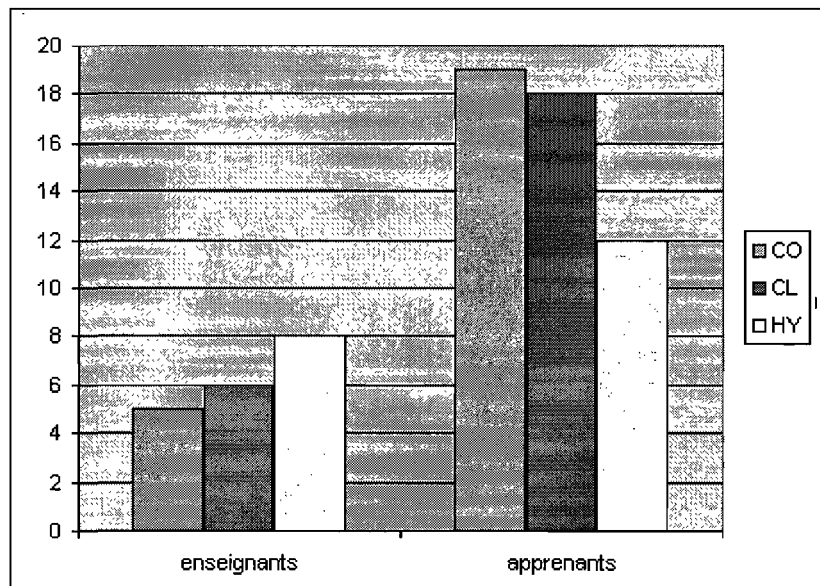


Figure 6.29 – Répartition des participants de chaque catégorie selon le profil

- Feedback des utilisateurs : Les utilisateurs sont invités à évaluer les recommandations qui leur sont fournies par le système de façon explicite et ce en introduisant les valeurs des explicites des pertinences pédagogiques.
- Domaine des ressources : Pour notre expérimentation, nous avons choisi des objets du domaine des sciences informatiques pour pouvoir utiliser l'ontologie que nous avons développée.
- Nombre de ressources évaluées : C'est le nombre de ressources dont dispose l'entrepôt d'objets d'apprentissage.

6.3.2 Méthode

Un utilisateur s'enregistre en remplissant le formulaire d'enregistrement du système, il est ensuite invité à se connecter. Il est alors aléatoirement affecté à l'une des trois catégories définies en 6.3.1 (voir tableau 6.2).

Selon la catégorie à laquelle il est affecté, il peut être soumis à l'un des trois cas de figures suivants :

- LORESA lui présente un ensemble de ressources que le recommandeur a sélectionné en se basant sur son profil. L'utilisateur est ensuite invité à demander les recommandations du système, à consulter les objets recommandés et enfin à évaluer ces recommandations.
- LORESA l'invite à chercher dans l'entrepôt les objets dont il a besoin. Et de la même façon que pour la première catégorie, il est invité à demander les recommandations du système, à consulter les objets recommandés et enfin à évaluer ces recommandations.
- LORESA lui présente un ensemble de ressources que le recommandeur a sélectionné en se basant sur son profil, le système lui demande ensuite s'il désire obtenir des recommandations concernant ces ressources. Dans ce cas, les recommandations sont générées et présentées à l'utilisateur qui peut les évaluer.

Avant de sortir du système, chaque utilisateur (quelque soit sa catégorie) est invité à répondre à un formulaire d'évaluation globale.

6.3.3 Validation des hypothèses

6.3.3.1 Utilisabilité du système

- *Les résultats* : Pour mesurer l'utilisabilité de notre système nous avons choisi de calculer le degré d'utilisation de chaque fonctionnalité du système selon le profil de l'utilisateur. La figure qui suit (figure 6.30) représente les résultats de cette étude dont les valeurs numériques sont données dans le tableau 6.4.

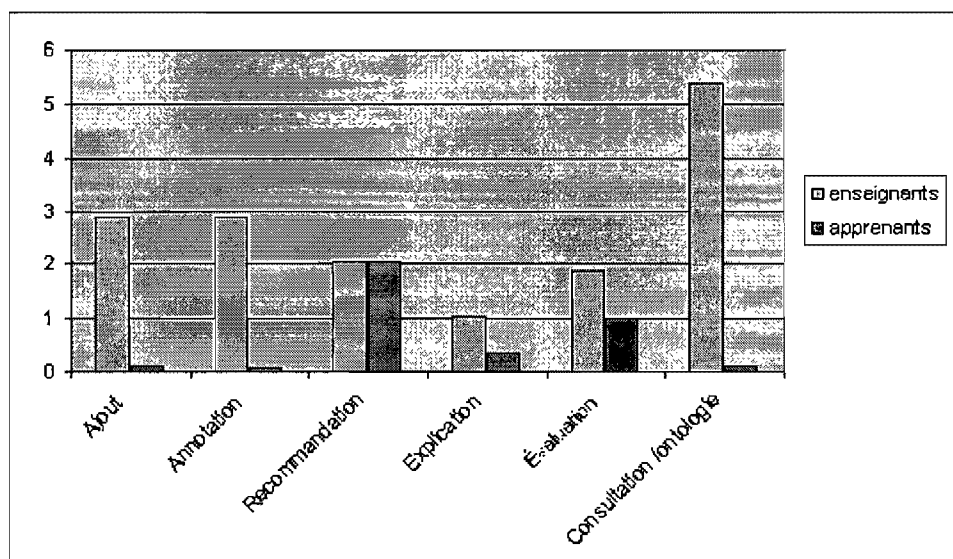


Figure 6.30 – Répartition des participants de chaque catégorie selon le profil

- *Interprétation des résultats* : Le premier critère que nous avons mesuré dans notre étude est le nombre de ressources ajoutées. En effet cette fonctionnalité qui permet d'enrichir l'entrepôt en ressources d'apprentissage est offerte aux deux catégories d'utilisateurs, cependant notons que cette fonctionnalité a été est de loin plus utilisée par les enseignants (nombre moyen de ressources ajoutées par utilisateur de 2,89) que par les apprenants (nombre moyen de ressources ajoutées par utilisateur de 0,10). Ceci nous conforte dans notre hypothèse selon laquelle la motivation première d'un apprenant pour la recherche de ressources est le besoin d'amélioration ou de renforcement de ses connaissances. Ce qui n'est pas le cas d'un enseignant qui cherche avant tout à partager les ressources qu'il a créées ou qu'il a utilisées avec d'autres utilisateurs quels que soient leurs profils. Dans le cas de la fonctionnalité de l'annotation, il est aussi évident que les participants du profil enseignant ont suivi le cheminement logique de la soumission d'une ressource à l'entrepôt, puisque le nombre moyen d'objets annotés (2,89) est égal nombre moyen d'objets ajoutés, c.-à-d. que toutes les ressources ajoutées ont été annotées. Notons que

Tableau 6.3 – Résultats de l'étude de l'utilisabilité de LORESA

	Enseignants	Apprenants
Nombre de participants	19	49
Nombre de ressources ajoutées	55	5
Nombre moyen de ressources ajoutées par utilisateur	2,89	0,10
Nombre de ressources annotées	55	3
Nombre moyen de ressources annotées par utilisateur	2,89	0,06
Nombre moyen de recommandations demandées	44	100
Nombre moyen de recommandations demandées par utilisateur	2,31	2,04
Nombre total d'explications demandées	20	34
Nombre moyen d'explications demandées par utilisateur	1,05	0,34
Nombre total des ressources évaluées	17	38
Nombre total d'évaluations	18	42
Nombre moyen d'évaluations par ressource	1,05	1,10
Nombre moyen d'évaluations par utilisateur	0,94	0,85
Nombre total des concepts utilisés pour l'annotation	102	6
Nombre moyen de concepts utilisés par utilisateur	5,36	0,12

ce n'est pas le cas des apprenants, où ne remarquons que le nombre moyen d'objets annotés (0,06) est inférieur au nombre moyen de ressources ajoutées (0,10). Ceci s'explique par le fait que les apprenants n'ont pas annoté toutes les ressources qu'ils ont ajoutées. Concernant les demandes de recommandations notons que le ratio utilisateur/recommandation est pratiquement le même pour les deux catégories d'utilisateurs (apprenants et enseignants), ceci s'explique par le fait que la recommandation est considérée par tout chercheur de ressource, comme un moyen d'aide

dans le processus de prise de décision. En effet, un utilisateur peut choisir de consulter la ressource qui est proposée car la valeur de sa recommandation est élevée. Le graphique (figure 6.30) nous démontre que la fonctionnalité la plus utilisée dans la catégorie des apprenants est la demande de recommandation, alors que dans la catégorie des enseignants c'est celle de l'ajout et l'annotation. Comme nous l'avons vu dans le chapitre implémentation, nous avons dans notre système une fonctionnalité qui permet d'obtenir à la demande l'explication de la recommandation obtenue. Notons que cette fonctionnalité a été utilisée à un degré plus élevé par la population des enseignants (1,05) que celle des apprenants(0,34). Une interprétation possible de cette donnée pourrait être la qualité de la recommandation fournie, qui était satisfaisante, donc qui ne nécessite pas une demande d'explications. Ce critère sera vu plus en détails dans la section suivante. Concernant l'évaluation des recommandations fournies, nous remarquons également que le ratio évaluation/utilisateur pour les deux populations n'est pas très éloigné avec tout de même un ratio plus élevé pour les enseignants (0,94 pour les enseignants et de 0,85 pour les apprenants). Ceci peut s'expliquer par le fait que l'enseignant est plus apte à juger de la pertinence pédagogique des ressources présentées. Le dernier critère que nous avons mesuré par le biais de cette étude est le degré d'utilisation de l'ontologie, car pour qu'un utilisateur puisse annoter son objet par le biais des concepts issus de l'ontologie, il doit d'abord la consulter. Étant donné que la fonctionnalité d'annotation a été plus utilisée par les enseignants que par les apprenants, ceci conduit systématiquement à une plus grande utilisation de la fonctionnalité de consultation de l'ontologie par cette même population.

6.3.3.2 Apport de la technique hybride du recommandeur

Comme nous l'avons expliqué en 6.3.1, nous voulons valider par cette étude la pertinence du choix de la technique de recommandation utilisée dans LORESA. En fait, cette dernière est une technique hybride d'un filtrage par contenu en cascade avec un filtrage

collaboratif. Pour ce faire nous allons déterminer le nombre des cas où l'utilisateur accepte la recommandation du système. Nous considérons qu'un utilisateur a accepté la recommandation s'il ne l'évalue pas.

- *La technique du filtrage par contenu* : Les résultats de la technique de filtrage basé sur le contenu sont donnés dans le tableau 6.4.

Nous remarquons de prime abord que le ratio recommandation demandée par usa-

Tableau 6.4 – Résultats de la technique du filtrage par contenu

	Ratio/usager	Demandée	Évaluée	Suivie	Pourcentage
Enseignants	2,8	14	9	5	35,71%
Apprenants	0,89	17	11	6	35,29%
Moyenne					35,50%

ger est plus élevé pour les enseignants que pour les apprenants. Ceci peut s'expliquer par le fait que les apprenants dans leurs processus de recherche, sont satisfaits par des ressources d'apprentissage conformes à leurs préférences telles que : format, langue, type, etc. Alors qu'un enseignant, même s'il préfère qu'une ressource soit au format qu'il préfère, doit avant tout, en juger de la pertinence. La technique du filtrage par contenu que nous avons appliquée pour ce cas se base pour la recommandation sur les préférences de l'utilisateur, les recommandations que reçoit l'utilisateur, sont uniquement basées sur les annotations sémantiques initiales des ressources d'apprentissage réalisées par leurs auteurs. Ceci explique à notre sens le nombre élevé des évaluations des recommandations. Nous pouvons dire que les chiffres de cette technique sont globalement peu satisfaisants puisque la moyenne des recommandations suivies pour les deux populations est de : 35,29%.

- *La technique du filtrage collaboratif* : Les résultats de l'application de cette technique, sont donnés dans le tableau 6.5. Dans le cas de cette technique, nous remarquons que le ratio de demande de recommandation des deux populations est rapproché. Dans ce cas, nous avons désactivé le filtrage par contenu, conduisant donc à la génération des recommandations basée sur la technique du filtrage col-

Tableau 6.5 – Résultats de la technique du filtrage collaboratif

	Ratio/usager	Demandée	Évaluée	Suivie	Pourcentage
Enseignants	1,16	7	3	4	57,14%
Apprenants	1,94	35	17	18	51,42%
Moyenne	1,55	21	10	11	54,28%

laboratif. Notons une amélioration notable du pourcentage des recommandations suivies pour les deux catégories avec une moyenne de : 54,28%. Nous pouvons dire que globalement, les résultats de l'application de cette technique sont satisfaisants.

- *La technique hybride* : Les résultats de l'application de la technique hybride basée sur le filtrage du contenu et le filtrage collaboratif, sont donnés dans le tableau 6.6. Nous remarquons que le ratio de demande de recommandation est le plus élevé

Tableau 6.6 – Résultats de la technique hybride

	Ratio/usager	Demandée	Évaluée	Suivie	Pourcentage
Enseignants	2,87	23	6	17	73,91%
Apprenants	4	48	14	34	70,83%
Moyenne	3,43	35,5	10	25,5	72,37%

pour les deux populations pour cette technique que pour n'importe quelle technique. Nous pouvons en déduire qu'étant donné que les deux techniques de recommandation sont appliquées, l'utilisateur a d'abord à sa disposition des ressources sélectionnées à partir de ces préférences, ensuite, le filtrage collaboratif est appliqué pour les prochaines recommandations.

- *La comparaison des trois techniques* : Les résultats des trois méthodes sont récapitulés dans le tableau 6.7

Ce tableau ainsi que le graphique (voir figure 6.31) font bien ressortir que la technique hybride, est celle avec le plus grand pourcentage de recommandations suivies. Ceci s'applique, que ce soit pour chaque population ou pour la moyenne des deux populations. Ceci nous conforte dans notre hypothèse du choix de la tech-

Tableau 6.7 – Résultats des trois techniques : CT, CO et HY

	CT	CO	HY
Enseignants	35,71%	57,14%	73,91%
Apprenants	35,29%	51,42%	70,83%
Moyenne	35,50%	54,28%	72,37%

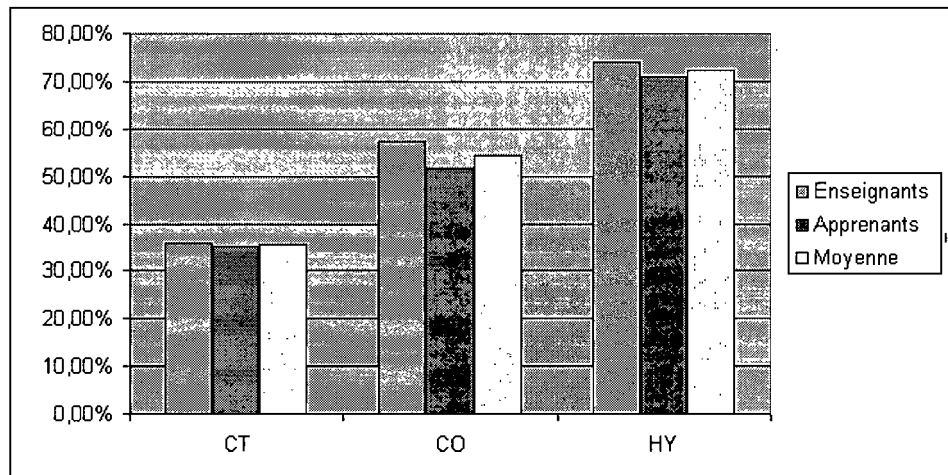


Figure 6.31 – Comparaison des trois techniques

nique hybride de recommandation. Nous pensons que l'avantage majeur de l'application de cette technique dans notre cas, est que même si l'utilisateur n'a pas fait de recherche, il se voit quand même présenter des ressources qui répondent à ses préférences, et demander ensuite des recommandations basées sur son profil.

6.3.3.3 Évaluation globale du système

Cette fonctionnalité est réalisée une seule fois par un utilisateur enregistré. Elle est accessible via le menu de navigation de la page principale, en cliquant sur le bouton *Évaluation*. À travers cette fonctionnalité, l'usager est invité à répondre à un formulaire d'évaluation globale du système et peut ajouter des commentaires au besoin (voir Figure 6.32). D'après le tableau 6.32, nous avons obtenu une moyenne de 4,7/5 concernant les objectifs du système, car nous pensons que nous avons expliqué dans la page d'accueil de notre système les objectifs de notre système et surtout, l'explication de chaque fonc-

Évaluation Globale du système LORESA

Bonjour, Jean

Afin d'améliorer notre système, nous vous remercions de bien vouloir répondre à ce formulaire. Dans toutes les questions, il vous est demandé d'attribuer une note entre 1 et 5 où 1 signifie faiblement d'accord et 5 totalement d'accord.

Les objectifs du système sont clairs : 1 2 3 4 5

L'utilisation de l'annotation sémantique améliore l'exploitation des objets d'apprentissage : 1 2 3 4 5

L'interface du système est conviviale : 1 2 3 4 5

L'explication de la recommandation est pertinente : 1 2 3 4 5

Les recommandations fournies sont précises : 1 2 3 4 5

Vos commentaires :

L'annotation sémantique aide à mieux indexer les ressources d'apprentissage.
Bonne continuation.

Figure 6.32 – Évaluation globale de LORESA

Tableau 6.8 – Résultats de l'évaluation globale de LORESA

	Moyenne des notes
Les objectifs du système sont clairs	4,7
L'utilisation de l'annotation sémantique améliore l'exploitation des objets d'apprentissage	4,5
L'interface du système est conviviale	4
L'explication de la recommandation est pertinente	3,9
Les recommandations fournies sont précises	4,1

tionnalité. Il en est de même pour la question concernant la pertinence de l'utilisation de l'annotation sémantique à des fins pédagogiques, où nous avons obtenu une moyenne de 4,5/5. Nous pensons que la fonctionnalité d'explication de la recommandation n'a pas été bien comprise par tous les utilisateurs, ceci s'est d'ailleurs confirmé par le faible ratio de demande d'explications par utilisateurs. Les deux autres critères d'évaluation ont été bien notés avec des moyennes de 4 et 4,1 respectivement. De plus, nous avons reçu des commentaires positifs et encourageants.

6.3.4 Conclusion

Nous avons par le biais de la validation, pu vérifier deux hypothèses que nous avons énoncées au départ. *La première* concerne l'utilisabilité du système, en effet nous avons supposé que les fonctionnalités de l'entrepôt d'apprentissage et aussi celles du système d'annotation soient plus utilisées par des usagers ayant un profil enseignant. Nous nous sommes basés dans cette hypothèse sur le fait qu'un enseignant est plus enclin à partager des ressources qu'il utilise ou qu'il crée avec une large communauté telle que les utilisateurs d'un entrepôt. L'apprenant, quant à lui, préfère utiliser un entrepôt pour rechercher des ressources dans un but de consolidation ou d'approfondissement de ses connaissances. L'étude que nous avons conduit, nous a conforté dans notre hypothèse.

Cette évaluation préliminaire nous a également permis de vérifier *la deuxième hypothèse*, à savoir le choix de la technique de recommandation : il s'est avéré que notre choix de cascader les deux techniques : contenu et collaboratif était le meilleur, car le pourcentage des recommandations suivies basées sur ces deux techniques est le plus élevé. Nous avons pu par le biais de cette étude, répondre à quelques interrogations concernant l'intérêt d'un tel système, la pertinence de l'annotation sémantique dans un contexte d'apprentissage, etc. Les résultats du questionnaire nous laissent penser que c'est le cas.

CHAPITRE 7

CONCLUSION

L'utilisation du Web dans des contextes éducationnels est une réalité que nous côtoyons de plus en plus de nos vies quotidiennes. Il devient nécessaire de disposer d'approches qui permettent aux enseignants et aux apprenants, de retrouver facilement et rapidement des ressources pédagogiques de qualité. Ces approches doivent également fournir un support au développement de communautés qui partagent les mêmes intérêts, pour améliorer leurs partages.

En partant de ce constat nous nous sommes fixés comme objectif de concevoir un système qui aide à recommander des ressources d'apprentissage à des usagers, en se basant sur des critères d'évaluation objectifs et formels. En effet, l'utilisation des techniques de filtrage collaboratif dans un contexte d'apprentissage n'est pas nouvelle, cependant, une attention particulière doit être portée au domaine d'application, car, le fait de l'utiliser directement dans le Web, conduit inévitablement à des résultats qui ne sont pas toujours pertinents, étant donnée l'hétérogénéité du contenu de ce médium de diffusion universel. Il était clair pour nous, que, pour arriver à recommander des ressources d'apprentissage de qualité, nous devons délimiter un espace restreint qui permettrait la création de communautés de partage. Nous avons alors pensé à un système de recommandation qui puisse recommander des ressources se trouvant dans un entrepôt d'objets d'apprentissage. La caractéristique de base d'une ressource dans un entrepôt d'objets d'apprentissage est sa description par des métadonnées. Un bref survol des entrepôts existants, nous a permis de conclure que quel que soit le schéma de description utilisé, la dimension pédagogique des ressources était complètement occultée.

Nous avons alors voulu combler cette lacune, en définissant un sous-schéma de description pédagogique de la ressource qui serait conjointement utilisé avec les métadonnées

classiques. Pour ce faire, nous avons défini cette description additionnelle sous forme *d'annotation sémantique*. Cette dernière permet de décrire de façon non ambiguë, puisqu'elle s'appuie dans cela sur une **ontologie du domaine**, les concepts couverts par la ressource. Le contenu de la ressource est évalué par celui qui le soumet à l'entrepôt, pour déterminer son apport à l'acquisition de ces concepts.

Il est inutile de rappeler l'importance du choix des critères d'évaluation d'un système de recommandation pour pouvoir générer des conseils de qualité. Ceci est d'autant plus vrai quand les articles recommandés sont des ressources d'apprentissage. Pour n'importe quel utilisateur, le plus important, dans son choix de ressource, demeure l'apport pédagogique de cette dernière. Il nous alors semblé tout naturel de baser les recommandations que nous voulions donner sur cette dimension importante.

À cet effet, nous avons conçu LORESA (Learning Objects REcommender system based on Semantic Annotations), comme nous l'avons décrit plus haut, l'inexistence dans les schémas actuels de description des ressources pédagogique, de la dimension pédagogique, nous a conduit à *définir* pour les ressources que nous voulions recommander, une *annotation sémantique*. Cette dernière de part sa définition, doit être basée sur une ontologie. Ceci a été le point de départ de notre conception, car si nous voulions ajouter à ces ressources des informations sur leur apport pédagogique, il était nécessaire de définir le vocabulaire que nous allons utiliser pour le faire. Nous avons, alors conçu notre ontologie dans le domaine des sciences informatiques, en se basant sur une classification existante que nous avons étendue pour pouvoir répondre à notre besoin. Nous sommes ensuite passé, à la création de l'entrepôt des ressources d'apprentissage pour que nous puissions annoter ces dernières par l'ontologie que nous avons définie.

Le principe de base de notre annotation sémantique des ressources, consiste à leur apporter l'information pédagogique manquante aux descriptions actuelles. Ainsi lorsqu'un utilisateur soumet un objet à l'entrepôt, il doit choisir trois concepts traités par son contenu. Ces trois concepts ne sont pas du texte libre, mais, des mots issus d'un vocabulaire

contrôlé en l'occurrence notre ontologie. Il évalue ensuite pour chaque concept, le degré de contribution de l'objet annoté à son acquisition. Ces informations sont utilisées comme entrée à notre système de recommandation.

Ce dernier s'appuie sur une technique hybride de recommandation, basé sur un filtrage par contenu et un filtrage collaboratif. La validation nous a conforté dans notre choix de la technique utilisée. LORESA applique cette technique de recommandation hybride afin de déterminer les objets d'apprentissage à recommander à un usager. Notre méthodologie a été vérifiée et les techniques employées ont montré qu'elles répondent à notre problématique (voir chapitre 6, section 6.3). En effet, le Tableau 7.7 montre que la recommandation fournie par la technique hybride a été suivie dans 72,37% des cas par l'utilisateur (contre 35,50% pour la technique du filtrage par contenu et 54,28% pour la technique collaboratif). De plus, les résultats obtenus dans la section 6.3.3.3 du chapitre 6, montrent que l'intérêt d'un tel système est démontré : l'idée originale et nouvelle d'utiliser des *annotations sémantiques* qui définissent la *pertinence pédagogique* de la ressource et utiliser cette information pour *la recommandation*.

Après avoir comparé notre système avec des systèmes de recommandations de ressources d'apprentissage RACOFI [Boley et al., 2003] et COLDEX [Baloian et al., 2004] (voir chapitre 5, section 5.5), LORESA s'est distingué par plusieurs points parmi lesquels : il est le seul système à utiliser les annotations sémantiques pour la recommandation des objets d'apprentissage et utilise à cet effet une ontologie intégrée. C'est également le seul qui donne un feedback à l'utilisateur sous forme d'explication de recommandations. Cette caractéristique apporte une aide considérable à l'utilisateur dans son processus de choix.

Concernant la problématique de la généralisation de méthodologie de LORESA, nous pouvons affirmer que notre système s'adapte à n'importe quel domaine : il suffit, pour

cela, de changer l'ontologie utilisée pour pouvoir annoter des ressources autres que dans le domaine de l'informatique : la biologie, la médecine, etc. Nous pouvons même l'appliquer à un autre domaine que des ressources pédagogiques, il suffirait pour cela de définir une autre sémantique de la pertinence pédagogique, qui serait par exemple un degré d'importance d'une caractéristique, pour juger de la qualité d'un produit. LORESA s'appuie pour la description pédagogique des ressources recommandées sur trois concepts, en effet, il peut s'avérer qu'un utilisateur qui désire soumettre un objet, trouve que le processus du choix de trois concepts et l'attribution des valeurs numériques est long. Cependant nous avons fait ce choix, car, il est rare dans le domaine de l'apprentissage, de trouver une ressource dont le contenu ne traite que d'un seul concept.

Étant donné que l'ontologie que nous avons développée est extensible, nous avons pensé étendre les fonctionnalités de LORESA, pour permettre par exemple à un utilisateur (ne disposant pas des privilèges de l'expert) qui désire annoter un objet, d'envoyer par courriel une demande d'ajout de concept à l'expert ou à l'administrateur. Ceci permettrait d'enrichir plus vite l'ontologie puisque cette tâche devient collaborative tout en garantissant la cohérence de cette dernière puisque l'ajout effectif ne peut être effectué que par les utilisateurs ayant les privilèges requis. Dans le futur, nous comptons également expérimenter LORESA sur plus d'utilisateurs.

BIBLIOGRAPHIE

- [ACM, 1998] The ACM Computer classification system.
<http://www.acm.org/class/1998/>, consulté en juin 2007.
- [Adomivicius et Tuzhilin, 2005] Adomivicius, G., Tuzhilin, A. *Toward the Next Generation of Recommender Systems : A Survey of the State-of-the-Art and Possible Extensions*. IEEE Transactions on knowledge and data engineering, vol. 17, No. 6, pp. 734-749, June, 2005.
- [Baloian et al., 2004] Baloian, N., Galdames, P., Collazos, CA., Guerrero, LA. *A Model for a Collaborative Recommender System for Multimedia Learning Material*. Groupware : Design, Implementation and Use, vol. 3198, pp. 281-288, November, 2004.
- [Berners-Lee et al., 1994] Berners-Lee, T., Cailliau, R., Luotonen, A., Frystyk, H., Secret A. The world-wide Web, *Communications of the ACM*, vol. 37, No. 8, pp. 76-82, August, 1994, .
- [Berners-Lee et al., 2001] Berners-Lee, T., Hendler, J., Lassila, O. The semantic Web. *The scientific American*, Vol. 31, May, 2001.
- [Boley et al., 2003] Boley, H., Tabet, S., Wagner, G. Design Rationale of RuleML : A Markup Language for Semantic Web Rules. In *Proceedings of Semantic Web Working Symposium (SWWS01)*. Stanford University, pp. 381-401 July, 2001.
- [Breese et al., 1998] Breese, J., Heckerman, D., Carl, K. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence*, pp. 43-52, 1998.
- [Breuker et al., 1999] Breuker, J. et Muntjewerff, A. Ontological Modelling for Designing Educational Systems. *Workshop on Ontologies for Intelligent Educational Systems, Ninth International Conference on Artificial Intelligence in Educa-*

- tion, AI-ED99, Le Mans, France, July 19-23, 1999. <http://www.ei.sanken.osaka-u.ac.jp/aied99/aied99-onto.html>
- [Bush, 1945] Bush, V. *As We May Think*, Atlantic Monthly. Disponible à <http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>
- [Burke, 2002] Burke, R. *Hybrid recommender systems : surveys and experiments*. User Modeling and User-Adapted Interaction, Vol.12 , No. 4, pp. 331-370, November, 2002.
- [Crampes et Ranwez, 2000] Crampes M., Ranwez, S. Ontology-supported and ontology-driven conceptual navigation on the World Wide Web. In *Proceedings of the 11th ACM on Hypertext and Hypermedia*. San Antonio, Texas, USA, pp. 191-199, June, 2000.
- [Claypool et al, 1999] Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., Sartin, M. Combining Content-Based and Collaborative Filters in an online Newspaper. In *Proceedings of AMC SIGIR Workshop on Recommender Systems*, Berkeley, CA, August 1999.
- [Desmontil et al., 2002] Desmontils, E., Jacquin, C. Annotations sur le Web : notes de lecture. *Journées de l'AS-CNRS Web sémantique*. Paris, France, 2002 disponible sur : <http://www.lalic.paris4.sorbonne.fr/stic/octobre/octobre1/Desmontils.pdf>
- [Denoue, 2001] Denoue, L. *De la création à la capitalisation des annotations dans un espace personnel d'informations* 135p. Thèse pour l'obtention du diplôme de doctorat en informatique : l'Université de Savoie, Octobre 2000, disponible sur : <http://www.fxpal.com/people/denoue/publications/theseoct2000.pdf>.
- [Devedzic, 1999] Devedzic, V. ITS Ontology Engineering : Borrowing from Design Patterns. *Workshop on Ontologies for Intelligent Educational Systems, Ninth International Conference on Artificial Intelligence in Education, AI-ED99*, Le Mans, France, July 19-23th, 1999. <http://www.ei.sanken.osaka-u.ac.jp/aied99/aied99-onto.html>

- [Downes, 2001] *Learning Objects : Resources For Distance Education Worldwide*. International Review of Research in Open and Distance Learning. Éducation@Canada, portail international de l'éducation au Canada. July, 2001.
- [Duitama et al, 2005] Duitama, JF., Defude, B., Bouzeghoub, A., Lecocq, C. *A Framework for the Generation of Adaptive Courses based on Semantic Metadata*. Multimedia Tools and Applications, Vol. 25, No. 3, pp. 377-390, March, 2005
- [Engelbart et al, 2005] Engelbart D., English W. A research center for augmenting human intellect. *AFIPS Conference Proceedings of the 1968 Fall Joint Computer Conference*, San Francisco, CA, vol. 33, p. 395-410, December, 1968.
- [Golbeck et al., 2002] Golbeck, J., Grove, B., Kalyanpur, A., Hendler, J. New tools for the semantic Web. *Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW 02)*. Sigüenza, Espagne. Lecture Notes in Artificial Intelligence, vol. 2473, Springer Verlag, p. 392-400, October, 2002.
- [Golberg et al., 1992] Goldberg, D., Nichols, D., Oki, B., Terry, D. *Using collaborative filtering to weave an information tapestry*. Communications of the ACM, vol. 35, No. 12, pp. 61-70, 1992.
- [Gruber, 1993] Gruber, Thomas R. *A Translation Approach to Portable Ontology Specifications*. Knowledge Acquisition, Vol. 5, No.2, pp. 199-220, 1993.
- [Grüninger et al., 1995] Grüninger, M., Fox, M. Methodology for the Design and Evaluation of Ontologies. *Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI, Montréal, 1995*
- [Guarino, 1997a] Guarino, N. Some Organizing Principles for a Unified Top-Level Ontology. *AAAI 1997 Spring Symposium on Ontological Engineering*, University of Stanford, California, 24 - 26 March, 1997. <http://www.ladseb.pd.cnr.it/infor/Ontology/Papers/OntologyPapers.html>

- [Guarino, 1997b] Guarino, N. *Semantic Matching : Formal Ontological Distinctions for Information Organization, Extraction, and Integration*. In *Information Extraction : A Multidisciplinary Approach to an Emerging Information Technology*, SCIE 1997, M. T. Paziienza (Eds.), Springer Verlag, pp. 139-170, 1997. <http://www.ladseb.pd.cnr.it/infor/ontology/Papers/OntologyPapers.html>
- [Habrant et al., 1999] Habrant, J., Corbel, A., Savoy J., *Utilisation des réseaux sémantiques pour la navigation dans l'hypertexte. Proceedings of Colloque Multimédia et Construction des Savoirs*, Mai, 1999.
- [Herlocker et al., 1999] Herlocker, J., Konstan, J., Borchers, A., Riedl, J. An algorithmic framework for performing collaborative filtering. *Proceedings of Research and Development in Information Retrieval*, 1999.
- [Ikeda et al., 1999] Ikeda, M., Hayashi, Y., Lai, J., Chen, W., Bourdeau, J., Seta, K., Mizoguchi, R. An ontology more than a shared vocabulary. *Workshop on Ontologies for Intelligent Educational Systems, Ninth International Conference on Artificial Intelligence in Education, AI-ED99*, Le Mans, France, July, 1999. <http://www.ei.sanken.osaka-u.ac.jp/aied99/aied99-onto.html>
- [Joubert et al., 1992] Joubert, M., Fieschi, M., Robert, J. *Représentation de concepts médicaux pour la recherche d'information : réalisation d'une maquette à partir de MMLS*. *Informatique et Santé , Nouvelles Méthodes de Traitement de l'Information en Médecine*, pp. 5-10, 1992.
- [Kahan et al., 2001] Kahan, J., Koivunen, M-R., Prud'Hommeaux, E., Swic, R. *Annotea : An Open RDF Infrastructure for Shared Web Annotations. WWW10, Tenth World Wide Web Conference*. Hong Kong, ACM Press, p. 626-632, May, 2001. disponible au : <http://www10.org/cdrom/papers/488/>
- [Karp, 1995] Karp, P.D. *A strategy for database interoperation*. *Journal of Computational Biology*, vol. 2, No. 4, pp. 573-586, 1995.

- [Kawazoe et al., 2003] Kawazoe, A., Collier, N. An ontologically-motivated annotation scheme for coreference. *International Workshop on Semantic Web Foundations and Application Technologies*. Nara, Japon, March, 2003. Available on : <http://www-kasm.nii.ac.jp/SWFAT/PAPERS/SWFAT14S.PDF>
- [Kiryakov et al., 2004] Kiryakov, A., Popov, B., Ognyanoff, D., Manov, D., Kirilov, A., Goranov, M. *Semantic Annotation*. Indexing, and Retrieval. Elsevier's Journal of Web Semantics, vol. 2, No. 1, December, 2004. <http://www.Websemanticsjournal.org/ps/pub/2005-10>
- [Konstantopoulos et al., 2001] Konstantopoulos, M., Darzentas, J.S, Koutsabasis, P., Spyrou, T., Darzentas, J. *Towards Integration of Learning Objects Metadata and Learner Profiles Design : Lessons Learnt from GESTALT*. Interactive Learning Environments, vol. 9, No. 3, pages 231-254, 2001.
- [Lafon et al., 2002] Lafon, Y., Bos B. Describing and retrieving photos using RDF and http. W3C Notes April, 2002. Available on : <http://www.w3.org/TR/photo-rdf/>. 19/07/2005.
- [Longmire, 2000] Longmire, W. A Primer on Learning Objects. *Learning Circuits*. <http://www.learningcircuits.org/mar2000/primer.html>
- [Mcgreal, 2004] Mcgreal, R. *Learning Objects : A Practical Definition* International Journal of Instructional technology and Distance Learning. vol. 1, No. 9, pp. 1-12, 2004.
- [Metros, 2005] Metros, S. Learning Objects : A Rose by another name . Educause, July/August 2005.
- [Mizoguchi, 1998] Mizoguchi, R. A step Towards Ontological Engineering. *The 12th National Conference on AI of JSAI*, pp. 24-31, June 1998. <http://www.ei.sanken.osaka-u.ac.jp/english/step-onteng.html>

- [Mizoguchi, 1998] Mizoguchi, R. Le rôle de l'ingénierie ontologique dans le domaine des eiah. *The 12th National Conference on AI of JSAI*, pp. 24-31, June, 1998. <http://www.ei.sanken.osaka-u.ac.jp/english/step-onteng.html>
- [Motta et al., 2000] Motta, E., Buckingham, S., Domingue, J. *Ontology-Driven Document Enrichment : Principles, Tools and Applications*. International Journal of Human-Computer Studies, 2000.
- [O'Sullivan et al., 2002] O'Sullivan, D., Wilson, D., Smyth, B. Improving case-based recommendation, a collaborative filtering approach. *In Proceedings of the 6th European Conference on Advances in Case-Base Reasoning, ECCBR 02*, Aberdeen, Scotland, Royaume-Uni, pp. 278-291, September, 2002
- [Paquette, 2004] Paquette, G. *Les Réseaux de banques d'objets d'apprentissage : potentiel et défis pour l'ingénierie pédagogique*. Communication présentée au colloque sur l'ingénierie pédagogique à l'heure des TIC. Novembre 2004. Montréal, Québec. <http://www.profetic.org:16080/colloque2004/article.php3>, consulté en octobre 2005.
- [Peccatte, 2002] Peccatte, JM. *Prolégomènes aux systèmes non-aristotéliens et à la sémantique générale de 1933 à 1950*. Traduction fr. Éditions de l'Éclat, pp. 7-64, 2002
- [Quillian, 1968] Quillian, R. *Semantic informatic processing*. Chapitre : Semantic memory, pp. 227-270, 1968.
- [Reeve et al., 2001] Reeve, L., Han, H. Survey of semantic annotation platforms. *Proceedings of the 2005 ACM symposium on Applied computing*. Santa Fe, New Mexico, USA, ACM Press, vol. 964, p. 1634-1638, March, 2005
- [Resnick et al., 1994] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J. (1994). GroupLens : An Open Architecture for Collaborative Filtering of Netnews. *In CSCW '94 : Conference on Computer Supported Cooperative Work*, ACM, pp. 175-186, 1994.

- [Resnick et al., 1997] Resnick, P., Varian, H. Recommender systems, Special Issue. *Communications of the ACM*, vol. 40, 1997.
- [Sarwar et al.,2001] Sarwar, B., Karypis, G., Konstan, J., Reidl, J. Item-based collaborative filtering recommendation algorithms. *In Proceedings of the 10th International Conference on World Wide Web*, Hong Kong, pp. 285-295, May, 2001,
- [Savia et al., 1998] Savia, E., Kurki, T., Jokela, S. Metadata Based Matching of Documents and User Profiles, *Proceedings in Human and Artificial Information Processing of the 8th Finnish Artificial Intelligence Conference*, . Jyväskylä, Finland, pp. 61-70, August, 1998.
- [Schoening, 1997] [Schoening, 1997] Schoening, R. A Case and Strategy for Developing Standardized Educational Domain Modules and Ontologies, <http://group-ieee.org/groups/ltsc/ontol.htm>, 1997
- [Shardanand et Patti, 1994] Shardanand, U., Patti, M. Social information filtering : Algorithms for automating "word of mouth". *In Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems*, pp. 210-217, 1995.
- [Sheth, 1994] Sheth, B. A Learning Approach to Personalised Information Filtering, M.I.T.<http://agents.www.media.mit.edu/groups/agents/papers/>.
- [Singh, 2000] Singh, H. Achieving Interoperability in e-Learning. *Learning Circuits*. <http://www.learningcircuits.org/mar2000/singh.html>
- [Smyth et Cotter, 1999] Smyth, B., Cotter, P. Surfing the digital wave : generating personalised television guides using collaborative, case-based recommendation. *In Proceedings of the 3rd International Conference on Case-based Reasoning*, Munich, Germany, pp. 561-571, July, 1999.
- [Sowa, 1984] Sowa, J. *Conceptual Structures : Information Processing in Mind and Machine*. Ed. Addison-Wesley, 1984.

- [Schwab, 2005] Schwab, D. Approche hybride - lexicale et thématique - pour la modélisation, la détection et l'exploitation des fonctions lexicales en vue de l'analyse sémantique de texte, 2005.
- [Texier, 2005] R. Texier. Taxinomies, thésaurus et ontologies. *EliKya, intelligence des organisations*, 2005.
- [Thesauri, 2001] Thesauri and controlled vocabularies. *National Library of Canada*, 2001.
- [Vasudevan et al., 1999] Vasudevan, V., Palmer, M. On Web Annotations : Promises and Pitfalls of Current Web Infrastructure. *Proceedings of the 32 th Hawaii International Conference on System Sciences*. Maui, Hawaii, January, 1999, IEEE Computer Society press. disponible au : <http://csdl.computer.org/comp/proceedings/hicss/1999/0001/02/00012012.PDF>
- [Wiley, 2001] Wiley, D. *The Instructional Use of Learning Objects*. Association for Instructional Technology. Association for Educational Communications and Technology. <http://www.reusability.org/read/> (consulté en octobre 2006)
- [Weinstein et al., 1997] Weinstein, P., Alloway, G. Seed Ontologies : growing digital libraries as distributed, intelligent systems. *Proceedings of the Second ACM Digital Library conference*, Philadelphia, PA, USA, July, 1997.
- [Weinstein, 1998] Weinstein, P. Ontology-Based Metadata : Transforming the MARC Legacy. *Actes Third ACM Digital Library conference*, Pittsburgh, PA, USA, June 1998.
- [Yan et Garcia-Molina, 1994] Yan, T.W., Garcia-Molina, H. Index Structures for Information Filtering under the Vector Space Model. *IEEE Conference on Data Engineering*, 1994