



Les systèmes d'information sémantiques et adaptatifs fondés sur des documents virtuels et le web sémantique

Serge Garlatti

► To cite this version:

Serge Garlatti. Les systèmes d'information sémantiques et adaptatifs fondés sur des documents virtuels et le web sémantique. Interface homme-machine [cs.HC]. IMT Atlantique; UBS, 2006. <tel-02462225>

HAL Id: tel-02462225

<https://hal.science/tel-02462225v1>

Submitted on 31 Jan 2020

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

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



HAL Authorization

Mémoire d'habilitation à Diriger les Recherches

Présenté devant

L'Université de Bretagne Sud

Serge Garlatti

Travail réalisé au sein

de l'Ecole Nationale Supérieure des Télécommunications de Bretagne

Sujet

Les systèmes d'information sémantiques et adaptatifs
fondés sur des documents virtuels
et le web sémantique

Soutenue le 23 Novembre 2006

Devant le jury composé de

Marc	Nanard	Rapporteur	LIRMM, Montpellier
Alain	Derycke	Rapporteur	TRIGONE, Lille
Jean	Charlet	Rapporteur	INSERM U739
Franck	Poirier	Examineur	UBS
Ioannis	Kanellos	Examineur	ENST Bretagne

Dédicace

Remerciements

Mes premiers remerciements vont aux les membres du Jury d'habilitation, pour leur confiance et pour avoir accepté d'évaluer mon travail. Mais je les remercie aussi pour toutes les interactions aussi bien scientifiques qu'humaines que j'entretiens avec eux, et cela depuis quelques années déjà.

Je suis très reconnaissant à Marc Nanard d'avoir accepté d'être rapporteur de ce travail et pour tous ses conseils avisés dans la préparation de cette habilitation. Je le remercie aussi pour toutes les interactions scientifiques que nous avons eues au cours de ces années, notamment au sein de la communauté française du document virtuel.

Alain Derycke a accepté d'être rapporteur et je le remercie pour la qualité de ces conseils et des échanges que nous avons développés depuis notre rencontre au sein du RTP 39, réseau thématique pluridisciplinaire « Apprentissage éducation et formation ».

Je remercie Jean Charlet d'avoir accepté d'être rapporteur. Je lui suis aussi très reconnaissant de m'avoir sollicité pour participer au RTP 33, réseau thématique pluridisciplinaire "Documents et contenu : création, indexation, navigation" et au comité éditorial de RFIA et pour tous les échanges que nous avons eus.

Je remercie Franck Poirier d'avoir accepté d'être membre de mon jury d'habilitation et pour ses conseils éclairés pour la préparation de cette dernière.

Ioannis Kanellos a accepté de participer à mon jury d'habilitation et je l'en remercie vivement. Mais notre collaboration et nos échanges remontent à bien plus longtemps comme nous allons le voir plus loin.

La suite de ces remerciements s'organise en deux temps et selon deux principes : 1) une organisation temporelle relative à des collègues et responsables de départements de l'ENST Bretagne, 2) une organisation par étape de recherche qui répartit ces remerciements à la fin des chapitres 1 à 4 de ce rapport dans le paragraphe « crédits ».

Je suis très reconnaissant à M. Jean Bézivin de m'avoir aidé et encouragé à faire un DEA informatique. Je le remercie aussi pour m'avoir encadré pour mon stage de DEA et ainsi me donner mon premier contact avec la recherche.

Ce premier contact s'est poursuivi par une inscription en thèse après quelques années passées dans l'industrie et à l'ENST Bretagne. Je remercie donc tout particulièrement Michel Habib qui a encadré ma thèse, m'a initié à la recherche et m'a laissé la liberté de construire ma problématique. Je le remercie aussi pour toutes les interactions que nous avons eues pendant de nombreuses années.

Après la thèse, Ioannis Kanellos et moi-même avons cette même volonté de développer et d'organiser la recherche au département informatique. La richesse de nos interactions, son soutien et nos diverses collaborations m'ont beaucoup apporté et je lui en suis très reconnaissant.

Je voudrais tout particulièrement remercier Jean Pierre Barthélemy – que Ioannis et moi-même avons sollicité pour qu’il développe la recherche au sein du département informatique. Cette recherche s’est ensuite poursuivie au sien d’un nouveau département appelé IASC, Intelligence Artificielle et Sciences Cognitives, dont il était le responsable. Jean Pierre Barthélemy a toujours su nous donner d’excellents conseils pour le développement de nos activités de recherche, mais aussi a su entretenir chez nous cette « fibre de la recherche ».

Je remercie aussi beaucoup Yvon Kermarrec pour toutes les fructueuses collaborations que nous avons et sans qui le projet CANDLE n’aurait pas eu existé.

Je voudrais également remercier Annie Gravey pour m’avoir accueilli - depuis peu - au sein du département Informatique et pour toutes les perspectives que ce changement va ouvrir au développement de mes problématiques de recherche.

Je remercie aussi toutes les personnes avec qui j’ai collaboré au cours de toutes ces années de recherche. La liste serait trop longue à faire, mais elles ont toutes contribué à m’amener au point où j’en suis aujourd’hui. Certaines d’entre elles sont présentes dans les paragraphes « crédits » des chapitres 1 à 4.

Table des matières

Mémoire d'habilitation à Diriger les Recherches	1
Chapitre 1. Introduction	19
1.1 Problématique de recherche	19
1.2 Plan du rapport et guide de lecture	23
Chapitre 2. Systèmes d'information sémantiques en médecine.....	27
2.1 Préambule.....	27
2.2 Introduction	29
2.3 Le projet ATLAS	35
2.3.1 Architecture du système d'information.....	36
2.3.1.1 L'hypermédia	37
2.3.1.1.1 L'assemblage ou présentation des différents nœuds	38
2.3.1.1.1.1 Les nœuds d'information	38
2.3.1.1.1.2 Les nœuds d'organisation.....	41
2.3.1.1.2 Le mécanisme de navigation	41
2.3.2 Architecture de la base de connaissances.....	42
2.3.2.1 Les connaissances de domaine	43
2.3.2.1.1 Les conflits d'héritage multiple	44
2.3.2.1.1.3 Le cadre général	44
2.3.2.1.2 Les différentes stratégies.....	46
2.3.2.1.2.4 La linéarisation.....	47
2.3.2.1.2.5 La désignation explicite	47
2.3.2.1.3 Les perspectives	49
2.3.2.2 Une base de connaissances intensionnelle	50
2.3.2.2.1 Nature des conflits.....	51
2.3.2.2.2 L'architecture de la base de connaissances.....	54

2.3.2.2.3	Le mécanisme d'héritage	55
2.3.2.2.4	Principes de l'algorithme de résolution de conflit	57
2.3.3	Conclusion.....	58
2.4	Le projet ATLAS-TUTOR.....	59
2.4.1	Problématique.....	59
2.4.2	Le projet MEDIANTE	60
2.4.3	Conception d'un ATLAS-TUTOR	63
2.4.3.1	L'instruction	64
2.4.3.2	L'indication	64
2.4.3.3	Remédiation	64
2.4.3.4	Interrogation	65
2.4.3.5	Gestion	66
2.4.3.6	Aide	66
2.4.4	Nouveaux types d'action.....	66
2.4.4.1	Applications d'un Atlas-Précepteur.....	67
2.4.5	Problèmes de conception.....	68
2.4.5.1	Un cadre unifié pour la connaissance.....	68
2.4.5.2	Enregistrement des planches d'atlas sur des images de patients	68
2.4.5.3	Raisonnement spatial.....	69
2.4.6	Conclusion.....	69
2.5	Conclusion générale	70
2.6	Crédits	70
2.7	Bibliographie.....	71
Chapitre 3.	Un système d'information adaptatif pour la navigation en mer.....	77
3.1	Préambule.....	77
3.2	Introduction	78

3.3	Compréhension d'un hypermédia	79
3.4	Les hypermédias adaptatifs	81
3.4.1	Les méthodes de présentation adaptative	82
3.4.2	Les méthodes de navigation adaptative.....	84
3.4.3	La modélisation utilisateur	86
3.4.3.1	L'acquisition des caractéristiques	86
3.4.3.2	Les types de modèles.....	87
3.4.3.3	Les caractéristiques de l'utilisateur.....	87
3.4.3.3.1	La connaissance.....	87
3.4.3.3.2	L'expérience, les compétences	88
3.4.3.3.3	Les préférences.....	88
3.4.3.3.4	Les objectifs	89
3.4.3.4	Les propriétés des caractéristiques utilisateurs	89
3.5	Le projet SWAN.....	90
3.5.1	Modélisation des objectifs et de l'organisation.....	92
3.5.1.1	Objectifs de recherche et de création de routes et d'information sur une route	93
3.5.2	Modèle utilisateur.....	95
3.5.3	Les catégories d'adaptation	96
3.5.3.1	Adaptation pour l'objectif de recherche et création de routes	97
3.5.3.2	Tâche « Affichage_Routes »	98
3.5.3.3	Adaptation pour la tâche de recherche d'information sur une route	99
3.5.3.4	Tâches « Aide à la Navigation »	101
3.5.3.5	Tâche « Informations IN »	103
3.5.4	Architecture du système	103
3.6	Conclusion.....	104
3.7	Crédits	105

3.8	Bibliographie.....	105
Chapitre 4. Les systèmes d'informations adaptatifs et sémantiques fondés sur les documents virtuels et le web sémantique..... 109		
4.1	Préambule.....	109
4.2	Introduction.....	111
4.3	Les documents virtuels.....	113
4.3.1	Historique.....	113
4.3.2	Définitions.....	113
4.3.3	Le processus de sélection.....	116
4.3.4	Le processus d'organisation.....	118
4.3.5	Le processus d'assemblage.....	120
4.3.6	Conclusion.....	121
4.4	Le web sémantique.....	121
4.4.1	L'architecture du web Sémantique.....	122
4.4.1.1	Le niveau adressage.....	123
4.4.1.2	Le niveau syntaxique.....	123
4.4.1.3	Le niveau sémantique.....	124
4.4.1.3.1	RDF (Resource Description Framework) et RDF Schéma.....	125
4.4.1.4	Les ontologies.....	126
4.4.1.5	Logique, Preuve et Confiance.....	128
4.4.2	Les métadonnées.....	128
4.4.2.1	Définitions.....	130
4.4.2.2	Scénario d'utilisation des métadonnées et de l'adaptation.....	132
4.4.2.2.1	Recherche d'information et création manuelle de cours.....	133
4.4.2.2.2	Composition dynamique de cours adaptatifs.....	134
4.4.2.2.3	Lecture et apprentissage par un apprenant.....	136
4.4.3	Conclusion.....	138

4.5	L'environnement SCARCE.....	138
4.5.1	Quelques principes pour la conception des systèmes d'information adaptatifs...	139
4.5.2	Spécifications et gestion séparées des composants des systèmes d'information..	140
4.5.3	Principes de conception.....	141
4.5.3.1	L'ontologie de document	143
4.5.3.2	L'ontologie des métadonnées.....	145
4.5.3.3	L'ontologie utilisateur.....	146
4.5.3.4	L'ontologie de domaine	147
4.5.3.5	Acquisition des connaissances pour l'ontologie de document.....	147
4.5.3.5.1.1	ICCARS	147
4.5.3.5.1.2	CANDLE.....	150
4.5.3.6	Le Moteur de composition adaptatif	157
4.5.3.7	La composition adaptative	159
4.5.3.8	Composition sémantique adaptative.....	159
4.5.3.9	Composition logique et composition de présentation	162
4.5.4	Conclusion.....	165
4.6	Crédits	167
4.7	Bibliographie.....	167
Chapitre 5.	Environnement informatique d'apprentissage humain	173
5.1	Préambule.....	173
5.2	Introduction	174
5.3	Problématiques et méthodologie	177
5.3.1	Principes fondamentaux de conception de l'EIAH	177
5.3.2	Questions de recherche.....	178
5.4	Méthodologie de co-conception des scénarii didactiques	180

5.5	Elaboration des scénarios d'apprentissage.....	181
5.5.1	Analyse des scénarios.....	183
5.5.2	Formalisation des scénarios par un modèle de tâches hiérarchiques	188
5.6	Rôle des scénarios dans la conception de l'EIAH.....	189
5.7	Conclusion.....	190
5.8	Crédits	191
5.9	Bibliographie.....	191
Chapitre 6.	Conclusion et perspectives	195
	Semim@ge	201
	SIAS	204
	P-LEARNET	205
	REACTIONS	206
	Objectifs	207
Chapitre 7.	Curriculum VITAE	209
7.1	Formations et qualifications	209
7.2	Expérience professionnelle	210
7.3	Activités d'enseignement	210
7.3.1	Enseignements dispensés (principalement à l'ENST Bretagne).....	210
7.3.2	Encadrements de projets.....	211
7.3.3	Ingénierie pédagogique	212
7.3.4	Activités d'organisation	212
7.3.5	Activités de recherche	213
7.3.6	Domaines et axes de recherche	213
7.3.6.1	Les enjeux	215
7.3.6.2	Quelques résultats	216
7.3.6.3	Quelques perspectives.....	217

7.3.7	Thèses encadrées ou dirigées	218
7.3.8	Stages de DEA encadrés ou dirigés.....	218
7.3.9	Administration et responsabilités collectives	219
7.3.10	Organisation de congrès	219
7.3.11	Participation à des comités scientifiques.....	220
7.3.12	Conférence invités, séminaires et tutoriaux	221
7.3.13	Mobilité, séjour Sabbatique.....	222
7.3.14	Participation à des jurys de thèse	222
7.3.15	Expertise de projet.....	223
7.3.16	Activités d'organisation et responsabilités internes à l'ENST Bretagne	223
7.4	Recherche contractuelle	223
7.4.1	Projet "Categories, Concepts and Symbolic Systems"	224
7.4.2	Projets ATLAS.....	224
7.4.3	Projet ATLAS-TUTOR.....	225
7.4.4	Projet SWAN.....	226
7.4.5	Projet ICCARS	227
7.4.6	Projet ERGOWEB.....	227
7.4.7	Projet CANDLE	228
7.4.8	Projet KMP.....	229
7.4.9	Projet MODALES	230
7.5	Publications	231
7.5.1	Revue	231
7.5.2	Contribution à des ouvrages collectifs	231
7.5.3	Actes de colloques internationaux avec comité de lecture.....	232
7.5.4	Colloques nationaux et colloques internationaux sans actes.....	233
Chapitre 8.	Bibliographie Générale	237

Table des figures

Figure 1: Processus d'utilisation d'un Atlas	31
Figure 2 : Exemple d'affichage en 3D.	32
Figure 3: Interprétation et appariement à partir d'un atlas.	33
Figure 4 : Organisation générale du système	37
Figure 5 : Type hiérarchie	38
Figure 6 : Type description	39
Figure 7 : Type texte	39
Figure 8 : Type Planche de Talairach.....	40
Figure 9 : Type illustration.....	40
Figure 10 : Nœud d'organisation	41
Figure 11 : Organisation de la base de connaissances dans ATLAS	43
Figure 12 : Conflits d'héritage multiple	46
Figure 13 : Exemples de points de vue dans deux systèmes.....	48
Figure 14: Relation de composition de type portion / masse en fonction des domaines	52
Figure 15: Relations de connexion entre substance-noire et striatum.....	53
Figure 16: Exemple de conflit inter domaine.....	53
Figure 17 : Organisation de la base de connaissances	55
Figure 18 : Exemple de graphe d'héritage avec conflits	56
Figure 19 : Interface de MR-TUTOR,	62
Figure 20 : Méthodes d'adaptation.....	82
Figure 21 : Décomposition d'un objectif recherche/création de routes	94
Figure 22 : Graphe d'héritage des tâches de recherche d'information.	95
Figure 23 : Modèle utilisateur	96
Figure 24: Vues spatiales dans SWAN.	100
Figure 25 : Modèle de domaine de SWAN	101

Figure 26 : Zone de pertinence.....	102
Figure 27 : Interface de SWAN.....	104
Figure 28 – Architecture du Web Sémantique	123
Figure 29 : Utilisation des métadonnées dans le web sémantique	130
Figure 30 : Modèle de document.....	144
Figure 31 : Poste de journalisme assisté par ordinateur.	150
Figure 32 : Système d'activité	152
Figure 33 : Structure d'un cours magistral ou d'une conférence	153
Figure 34 : Structure d'un « conflit ».....	153
Figure 35: Structure d'un processus de discussion	154
Figure 36 : vue globale de l'outil CAT (CANDLE Authoring Tool)	155
Figure 37 : structure d'une activité avec le CAT	156
Figure 38 : Saisie des métadonnées dans CAT	156
Figure 39 : Le cours CANDLE sur ATM	157
Figure 40 : L'architecture de moteur de composition.....	157
Figure 41 : Un document générique.....	158
Figure 42 : Disposition de page Web	159
Figure 43 : La sélection.....	160
Figure 44 : Évaluation : règles d'appartenance aux classes	161
Figure 45 : Choix d'une méthode d'adaptation.....	161
Figure 46 : Application d'une technique adaptative	161
Figure 47 : paramètre reçu par le moteur de servlet.....	162
Figure 48 : Guide local de navigation : composant XML.....	163
Figure 49 : Une organisation des étapes du scénario P0.....	185
Figure 50 : Schéma d'un système tâche/technique routinier	186
Figure 51: Schéma d'un système tâche/technique en apprentissage	187

Chapitre 1. Introduction

1.1 Problématique de recherche

Dans un premier temps, l'activité de recherche était principalement orientée vers la représentation de connaissances et le raisonnement (représentations objets, héritage multiple avec exceptions, raisonnement et logique non monotone). Cette problématique était déjà orientée vers la prise en compte du facteur humain. En effet, il s'agissait de proposer des mécanismes d'héritage avec exceptions qui avait du sens pour l'homme. Plus tard, les relations de cette thématique avec certaines théories cognitives de la catégorisation (théorie du prototype) ont été étudiées dans une perspective pluridisciplinaire au sein du GDR de sciences cognitives (GDR 957).

Dans un second temps, cette première thématique a été étendue vers la conception de systèmes d'information sémantiques et adaptatifs. De nos jours, les systèmes d'information (SI) sont conçus pour être accessibles sur Internet (hypermédia), possèdent une architecture distribuée, accèdent à une masse de ressources hétérogènes qu'il s'agit de réutiliser, assurent une adaptation / personnalisation en fonction des besoins utilisateur et permettent la coopération homme/machine et/ou homme/homme médiatisée par des machines au travers du web. Un système d'information adaptatif peut proposer différents contenus, différentes organisations et différentes présentations en fonction des besoins utilisateur. Ces systèmes d'information sont appelés sémantiques parce que la description des contenus, de l'organisation et de l'adaptation est fondée sur une modélisation du sens par des

représentations de connaissances formelles. La conception de ces systèmes pose les problèmes suivants :

- Association du sens aux ressources
 - on peut imaginer que plus on décrit le sens à des niveaux sémantiques à la fois profonds et étendus plus on pourra améliorer la qualité de la recherche d'information, et donc plus on pourra automatiser certaines tâches. Il est bien évident qu'il y a une sorte d'optimum et/ou de ratio coût/gain acceptable dans chaque situation qui fixe les limites de la description du sens. Il y a malgré tout une limite due aux types d'activités mis en œuvre : certaines activités nécessitant une réelle expertise humaine ne sont pas transposables en machine. Il n'est pas possible d'acquérir les connaissances correspondantes.
- La répartition de l'interprétation et de l'action entre l'homme et la machine
 - Elle est directement liée à celle de l'attribution du sens. Dans une certaine mesure, fixer les limites de l'attribution du sens, c'est aussi fixer la répartition de l'interprétation et de l'action puisque les descriptions sémantiques ont pour objet de permettre à la machine d'automatiser et/ou de semi automatiser certaines activités.
- La compréhension et la cohérence du système pour assurer la coopération homme/machine :
 - En science cognitive, la compréhension d'un document hypermédia est souvent caractérisée par la construction mentale d'une représentation, d'un modèle de ce document. La lisibilité du document peut être définie comme l'effort mental nécessaire au processus de construction d'un modèle. Afin d'accroître la lisibilité d'un document hypermédia, il est nécessaire d'assister le lecteur dans la construction de ce modèle mental, par exemple en favorisant les facteurs utiles pour le processus de construction et en diminuant les facteurs qui le perturbent. Deux facteurs principaux ont été identifiés : la cohérence et le surcoût cognitif. Pour augmenter la cohérence d'un document, il faut fournir au lecteur des repères l'aidant à identifier les composants majeurs du document et son organisation. Pour réduire l'effort mental de compréhension, il n'est pas suffisant d'avoir une organisation cohérente, il faut aussi communiquer cette dernière au lecteur – c'est l'un des rôles de la présentation. Il est alors possible de fournir au lecteur les différents composants du document ainsi que leurs relations. Pour assurer cette cohérence, il est nécessaire que le contenu, l'organisation et la présentation aient une sémantique reconnue par l'utilisateur.
- Organisation et acquisition des connaissances
 - Comme le SI est fondé sur des connaissances, il est nécessaire d'avoir une méthodologie d'acquisition et de structuration des connaissances qui permettent une acquisition et une mise à jour aisée de celles-ci. En d'autres termes, il faut déterminer les différentes catégories de connaissances et leurs rôles dans le SI.

- La recherche d'information
 - Une augmentation de la complexité liée à la masse des ressources disponibles et leur hétérogénéité pose des problèmes de recherche d'information mais aussi de réutilisation, de partage et d'échange. La recherche d'information par mots clés sur Internet possède de fortes limitations qui ne permettent pas de satisfaire les besoins nécessaires à la conception de ces systèmes. Il faudra donc proposer de nouvelles méthodes de recherche d'information.
- Les méthodes de conception
 - Il est nécessaire de proposer de nouvelles méthodes facilitant la maintenance et la conception de ces systèmes – par exemple, un système d'information par utilisateur n'est pas envisageable. En d'autres termes, il est nécessaire d'automatiser le plus possible la création de ces systèmes, notamment pour les adapter aux besoins utilisateur.

Les quatre premières problématiques reflètent certains aspects de la prise en compte du facteur humain dans la conception de ces systèmes d'informations. Nous verrons tout au long de ce rapport que la gestion du contexte et de l'adaptation en sont d'autres manifestations. L'objet principal de cette recherche est de définir une classe de méthodologies adaptées à la conception de systèmes d'information sémantiques et adaptatifs assurant une distribution du rôle de l'interprétation et de l'action entre l'homme et la machine¹. L'approche proposée est fondée sur :

- Des hypermédias adaptatifs qui fournissent un cadre conceptuel intéressant pour l'adaptation des systèmes d'information. Un hypermédia adaptatif est un hypertexte ou hypermédia qui adapte le contenu, la navigation et la présentation en fonction d'un modèle utilisateur. Les modèles utilisateur, issus de l'intelligence artificielle, prennent notamment en compte les connaissances, les tâches, les préférences et les compétences d'un utilisateur ou d'un groupe d'utilisateurs.
- Le web sémantique et ses standards et la représentation des connaissances : association de « sémantique » à l'adaptation, à l'organisation et aux ressources pour améliorer la recherche d'information et permettre la réutilisation de ressources et une automatisation de la création des systèmes d'information.
La recherche d'information par mots clés sur Internet possède de fortes limitations en termes de précision et de rappel. Par contre, les moteurs de recherche sémantique, fondés sur l'initiative du web sémantique - des métadonnées sémantiques, les ontologies correspondantes et des moteurs d'inférence - surmontent ces limitations en associant de la sémantique aux ressources. Les métadonnées et les ontologies correspondantes se substituent aux ressources et définissent un contexte de

¹ Parler d'interprétation pour une machine est un terme un peu abusif. On fait uniquement l'hypothèse que la machine exécute un algorithme équivalent à une démonstration automatique de théorèmes, que le système formel correspondant a les propriétés de complétude et de correction et donc que toutes les opérations faites par cette machine (la démonstration de théorèmes) respecte la sémantique de la théorie des modèles correspondante (tous les théorèmes sont des formules valides).

réutilisation de celles-ci. Le web sémantique améliore la qualité de la recherche d'information et autorise donc le partage et la réutilisation des ressources. Il permet ainsi de rendre le contenu du web interprétable par les machines. Il devient donc possible d'automatiser et/ou de semi automatiser la création des SI.

- Des usages qui déterminent un contexte d'interprétation et d'action dans une communauté de pratiques particulière.
L'approche proposée s'appuie sur des pratiques : il s'agit de composer un document réel pour un utilisateur ayant des objectifs précis dans un contexte particulier : une communauté de pratiques. Les connaissances, compétences et savoir-faire explicites issus de ces communautés de pratiques permettent de donner du sens et d'assurer la cohérence des résultats.
- Les documents virtuels : méthodologie permettant l'automatisation de la création de SI et facilitant la maintenance de systèmes d'information réutilisant des ressources.
Un système d'information sémantique et adaptatif peut être vu comme un document virtuel adaptatif. Les documents virtuels sont des documents qui n'ont pas d'état persistant et pour lesquels chaque instance – document réel - est généré dynamiquement en fonction de la demande utilisateur. En d'autres termes, un document virtuel adaptatif possède un moteur de composition qui sélectionne, organise, adapte et assemble dynamiquement des ressources multimédia pertinentes. Le document réel est généré dynamiquement à partir de modèles sémantiques – ontologies – et d'une spécification. Une spécification est un paramètre du moteur de composition qui peut être capable de définir au plus les critères de sélection, d'organisation, d'adaptation et d'assemblage - présentation. Les documents virtuels apportent des méthodes et des techniques qui facilitent la création et la maintenance des systèmes d'information.

En résumé, ces systèmes d'information sémantiques et adaptatifs sont fondés sur des documents virtuels adaptatifs dont :

- Les connaissances explicites et partagées de la communauté de pratiques sont représentées par des ontologies. Pour assurer la cohérence, le contenu, l'organisation et l'adaptation sont fondés sur ces connaissances. Ils devraient ainsi avoir une sémantique reconnue par l'utilisateur.
- La sélection utilise un schéma de métadonnées ou un mécanisme d'indexation, les ontologies et un moteur d'inférence pour retrouver les ressources pertinentes à partir des critères de sélection fournis dans la spécification.
- La spécification de la sélection, de l'adaptation et de l'organisation est faite à un niveau sémantique en utilisant un schéma de métadonnées ou une indexation et des ontologies. La spécification a aussi pour rôle de lier entre eux la sélection, l'adaptation et l'organisation afin de garantir la cohérence du résultat.

1.2 Plan du rapport et guide de lecture

Ce rapport s'organise autour de six chapitres :

- Les quatre premiers retracent les quatre étapes significatives de l'activité de recherche. Ces quatre étapes sont organisées autour de projets concrets avec une approche pluridisciplinaire définissant un contexte d'interprétation et d'action dans une communauté de pratiques particulière. Dans chaque étape, le ou les projets concernés partagent des problématiques, des méthodes et des technologies communes qui retracent les évolutions significatives de l'activité de recherche. Chacun de ces chapitres commence par un préambule qui définit les propriétés principales de l'étape concernée, issues des problèmes et enjeux principaux cités dans cette introduction. Chacun de ces chapitres contient un paragraphe « Crédits » qui précise les différentes contributions à cette activité (concept emprunté à J. Charlet).
- Le cinquième est constitué par une conclusion suivi des perspectives
- Le dernier appelé « Curriculum Vitae » propose une vision synthétique des activités d'enseignement, de recherche, des résultats et des publications. Il permet notamment d'avoir un « résumé » des quatre premiers chapitres.

Les références bibliographiques sont proposées par chapitre pour les quatre étapes de l'activité de recherche – les quatre premiers chapitres. Chacun d'entre eux regroupant toutes les références propres à chaque étape. Le chapitre « CV » propose toutes mes publications classées. Le dernier chapitre reprend toutes les bibliographies des chapitres 1 à 5.

Nous allons maintenant donner une vue plus détaillée des quatre étapes principales retraçant l'évolution de l'activité de recherche et des technologies au travers des projets en explicitant leurs caractéristiques communes pour chaque étape et les différences entre les étapes :

- **Chapitre 2 :** La première étape vise à concevoir des systèmes d'information sémantiques pour l'aide au diagnostic dans le domaine médical. Les problématiques principales de recherche abordées sont : i) Recherche d'information et d'actions pertinentes en fonction de la situation courante de diagnostic, ii) Représentation de connaissances et raisonnement. L'aide au diagnostic nécessite d'articuler des connaissances générales (médicales) et des connaissances en situation (données patient) afin d'aider l'utilisateur dans sa tâche d'expertise - diagnostic. Les connaissances générales et explicites de la communauté de pratiques considérée permettent d'organiser l'accès aux informations et actions et de rechercher celles qui sont pertinentes en fonction de la situation courante. Il est donc nécessaire de déterminer les différentes catégories de connaissances, leurs rôles et leur organisation relative pour favoriser cette articulation. Dans un tel contexte, la répartition de l'action et de l'interprétation se fait de la manière suivante : le processus de « résolution de problème » est entièrement contrôlé par l'utilisateur – l'expert puisqu'il s'agit d'un processus heuristique et progressif qui peut exiger des étapes de retours arrière fréquents. Le système d'information ne peut donc que proposer des mécanismes de recherche et d'accès aux informations, actions, aux connaissances pertinentes et aux outils de visualisation et de traitement d'images relativement à la tâche ou au domaine

de l'utilisateur en garantissant la cohérence entre toutes les sources d'information et de connaissances. Les tâches/objectifs des utilisateurs sont pris en compte de manière implicite - il n'y a ni modèle utilisateur, ni adaptation, par contre les informations et/ou les connaissances sont proposées en fonction de la situation courante : le domaine d'expertise pour ATLAS et les erreurs détectées pour ATLAS-TUTOR. La notion de document virtuel n'est pas encore présente, par contre le système sélectionne les informations, les actions et les connaissances pertinentes et génère dynamiquement un hypermédia qui permet l'accès à celles-ci. Les informations et les connaissances ne sont réutilisées que localement. Il n'y a pas de partage et de réutilisation à distance de celles-ci. Pour le projet ATLAS, il s'agit de concevoir un système d'aide au diagnostic pour le traitement des épilepsies focales, différents modèles de domaine du cerveau (neuroanatomie, neurophysiologie, etc.). Ce système ne doit fournir à l'utilisateur que les informations pertinentes en fonction du contexte – domaine d'expertise de l'utilisateur. Pour le projet ATLAS-TUTOR, il s'agit de concevoir un tuteur intelligent d'aide au diagnostic pour la détection de lésion du cerveau –. Ce projet a notamment étudié l'apport du raisonnement spatial pour la détection d'erreurs et la remédiation.

- **Chapitre 3** : la seconde étape se distingue de la précédente par la nature des tâches utilisateur et par la prise en compte de l'adaptation aux besoins utilisateur. L'adaptation aux besoins utilisateur est un raffinement de la notion de situation courante de l'étape précédente. Il ne s'agit plus de tâches d'expertise pour les utilisateurs. Il est donc possible d'introduire un peu plus d'automatisation dans le SI. La problématique principale devient donc celle de la conception d'un système d'information sémantique et adaptatif. Ce dernier nécessite notamment la conception d'un modèle utilisateur et d'un mécanisme d'adaptation. Le système d'information doit prendre en charge une part plus importante : des mécanismes de recherche d'information, une structure organisant l'accès aux informations pertinentes en fonction des besoins utilisateur et des mécanismes d'adaptation – filtrage des informations pertinentes. La méthodologie de conception d'un tel système d'information doit donc être revue, notamment par la nécessité d'utiliser de nouvelles catégories de connaissances. Le système d'information est ici accessible sur internet au travers d'un navigateur. Les informations ne sont réutilisables que localement. Cette étape s'est déroulée autour du projet SWAN. Ce dernier a pour objet la conception d'un système d'aide à la navigation en ligne pour des marins. Les connaissances explicites de la communauté de pratiques concernée sont constituées d'un modèle de domaine qui permet d'indexer les informations, d'un modèle de tâches hiérarchiques décrivant les tâches utilisateur et organisant l'accès aux informations, d'un modèle utilisateur, d'un modèle de contexte, et d'un modèle d'adaptation. La recherche d'information et les requêtes associées pour la sélection sont de même nature qu'à l'étape précédente. Le système d'information adaptatif est engendré dynamiquement à partir des modèles précédents. Le concept de document virtuel commence à apparaître, mais n'est pas encore complètement exploité. C'est le modèle de tâche hiérarchique qui permet de spécifier l'organisation.
- **Chapitre 4** : Cette troisième étape marque une transition importante tant du point de vue conceptuel que technologique. En effet, les informations utilisées pour engendrer le système sont maintenant accessibles et distribuées sur internet. Il est donc nécessaire d'assurer leur recherche, leur réutilisation et leur partage par différents

systèmes d'information. Cette réutilisation pose les problèmes suivants : une augmentation de la complexité liée à la masse d'informations disponibles et à leur hétérogénéité, la recherche d'information par mots clés sur Internet possède de fortes limitations en termes de précision et de rappel et n'est donc plus adaptée à une réutilisation par des machines de ces informations, l'interopérabilité entre les informations et les systèmes qui les réutilisent, l'automatisation ou la semi automatisation de certaines tâches - actions - du système qui doivent par exemple organiser les informations de manière cohérente - ce problème est étroitement lié à la qualité de la recherche d'information puisqu'il s'agit notamment d'obtenir les informations pertinentes au bon moment et à la bonne place dans l'organisation.

Le concept de document virtuel dont le moteur de composition permet la réutilisation d'information et donc une forme d'automatisation devient fondamental. Il en est de même pour le web sémantique muni de ses standards – W3C. En effet, les moteurs de recherche sémantique constitués de métadonnées sémantiques, des ontologies correspondantes et des moteurs d'inférence, surmontent les limitations de la recherche par mots clé en associant du sens aux ressources - informations. Les métadonnées et les ontologies correspondantes se substituent aux ressources et définissent un contexte de réutilisation de celles-ci afin d'améliorer la qualité de la recherche d'information. Un moteur de composition doit donc sélectionner, adapter, organiser et assembler des ressources hétérogènes, distribuées et décrites par des métadonnées sémantiques et des ontologies.

Dans cette troisième étape, on peut notamment citer les enjeux suivants :

- La description du sens, la distribution de l'interprétation et de l'action entre l'homme et la machine et l'interopérabilité des ressources au niveau sémantique :

Ces deux enjeux ont plus d'importance qu'aux deux premières étapes parce que les ressources, leurs descriptions par des métadonnées sémantiques et les ontologies correspondantes ne sont plus maîtrisées localement, mais peuvent être produites par une communauté géographiquement répartie. La description du sens est réalisée à partir d'un schéma de métadonnées et des ontologies correspondantes qui assurent l'interopérabilité au niveau sémantique. Il est donc d'autant plus nécessaire de coordonner et de partager les connaissances de cette communauté, les guides de bonnes pratiques et la méthodologie pour assurer une bonne qualité de cette description.

- La méthodologie de conception de ces systèmes d'information sémantiques et adaptatifs doit donc être complètement modifiée pour assurer le partage et la réutilisation des ressources et l'adaptation. Dès que l'on dispose du concept de moteur de composition permettant d'engendrer dynamiquement un système d'information, on est amené à se poser la question de sa réutilisation dans des contextes différents. Serait-il possible de réutiliser ce moteur de composition pour d'autres types de systèmes d'information ? En d'autres termes, comment assurer une bonne flexibilité et extensibilité du moteur de composition ?

Flexibilité : plus les mécanismes du moteur de composition sont paramétrables, plus il est facile de maintenir et de concevoir de nouveaux systèmes d'information. De nouveaux documents réels peuvent être générés dès que de nouveaux paramètres de la sélection, de l'adaptation, de l'organisation et de l'assemblage

sont définis et/ou modifiés.

Extensibilité : à tout document virtuel sont associées des propriétés fondamentales qui caractérisent le type des systèmes d'information produits. Un moteur de composition met en œuvre ces caractéristiques. Un moteur de composition n'est réutilisable que si ces propriétés fondamentales sont adéquates aux domaines d'application envisagés. En d'autres mots, si ces principes sont compatibles avec la conception d'un type de système d'information particulier, on peut le réutiliser. Il est bien évident que ces principes fondamentaux sont aussi une limitation à la réutilisation. La conception du moteur de composition doit donc faciliter l'extension de ces propriétés.

Cette troisième étape a été constituée des projets suivants :

- Projet ICCARS : Composition sémantique et adaptative de dossiers thématiques et projet ERGO-WEB pour l'évaluation de l'expertise des journalistes et du poste de travail proposé.
- Projet CANDLE : Composition sémantique et adaptative de cours en ligne
- Projet KMP : Portail sémantique pour du « Knowledge Management »
- **Chapitre 5** : Cette quatrième étape est la suite de l'étape précédente puisqu'elle en possède presque toutes les caractéristiques. En effet elle a pour objet la composition sémantique et adaptative de cours en ligne fondée sur des pratiques et des théories didactiques. Mais par rapport à l'étape précédente, l'acquisition des connaissances explicites de la communauté de pratiques est plus difficile. En effet, l'introduction de nouveaux environnements informatiques d'apprentissage humain ou EIAH va profondément modifier les pratiques courantes. Il va donc être nécessaire de mettre en œuvre une méthodologie d'acquisition de ces connaissances adaptée à ce contexte. Cette dernière doit aussi s'intéresser à l'évaluation des résultats pour valider et/ou modifier les pratiques engendrées. Cette étape s'inscrit dans le cadre du projet MODALES (MOdelling Didactic-based Active Learning Environment in Sciences). Elle concerne plus particulièrement la mise en place de séquences de formation, à destination de futurs enseignants, (appelés ci-après apprenants) fondées sur l'utilisation de ressources multimédias et s'appuyant sur des savoir-faire d'experts en éducation et sur la modélisation de pratiques réelles. Il s'agit de mettre en place des scénarios didactiques qui permettent à de futurs enseignants « d'apprendre leur métier » en réalisant des séquences d'enseignement à partir d'une approche de type pédagogie active. L'attribution de l'interprétation et de l'action entre les apprenants, l'enseignant et la machine se fera dynamiquement en fonction des apprenants et de la situation courante. La conception d'un EIAH nécessite de mettre en œuvre des théories didactiques, des théories de l'activité humaine, ainsi que des modèles et théories informatiques. Se pose alors la question de l'articulation entre la modélisation des situations et leur transposition informatique. Ce projet en cours de réalisation, nous a amené à mettre en œuvre une méthodologie de co-conception de scénarii didactiques fondée sur des pratiques réelles et une approche transdisciplinaire comme fondement d'un EIAH.

Chapitre 2. Systèmes d'information sémantiques en médecine

2.1 Préambule

Cette première étape a pour objet la conception de systèmes d'information sémantiques d'aide au diagnostic dans le domaine médical. Les problèmes de recherche principaux sont la recherche d'information et d'actions pertinentes en fonction de la situation courante de diagnostic, la représentation de connaissances et le raisonnement. Dans ce contexte, le diagnostic nécessite d'articuler des connaissances générales (médicales) et des connaissances en situation afin d'aider l'utilisateur dans sa tâche d'expertise. Les connaissances générales et explicites de la communauté de pratiques considérée permettent d'organiser l'accès aux informations et actions et de rechercher celles qui sont pertinentes en fonction de la situation courante. Il est donc nécessaire de déterminer les différentes catégories de connaissances, leurs rôles et leur organisation relative pour favoriser cette articulation. La répartition de l'action et de l'interprétation se fait de la manière suivante : le processus de « résolution de problème » est entièrement contrôlé par l'utilisateur – l'expert puisqu'il s'agit d'un processus heuristique et progressif qui peut exiger des étapes de retours arrière fréquents. Le système d'information ne peut donc que proposer des mécanismes de recherche et d'accès aux informations, actions, connaissances pertinentes et aux outils de visualisation et de traitement d'images relativement à la tâche ou au domaine de l'utilisateur en garantissant la cohérence entre toutes les sources d'information et de connaissances. Les tâches/objectifs des utilisateurs sont pris en compte de manière implicite - il n'y a ni modèle utilisateur, ni adaptation, par

contre les informations et/ou les connaissances sont proposées en fonction de la situation courante : le domaine d'expertise pour ATLAS et les erreurs détectées pour ATLAS-TUTOR.

Pour ATLAS, le système d'information est fondé sur un hypermédia. Cela permet de fournir à l'utilisateur des informations pertinentes en fonction du contexte. Nous avons retenu les critères suivants pour le contexte : discipline médicale concernée (par exemple, neurophysiologie, neuroanatomie, neurochimie) et tâche à réaliser. Seuls les concepts pertinents dans une discipline sont décrits dans celle-ci, cela permet d'une part de fournir à l'utilisateur parmi différentes descriptions possibles d'un concept relatif au cerveau, celle qui est adaptée au contexte et d'autre part cela permet de restreindre les informations accessibles en fonction de la discipline considérée afin de conserver la cohérence des informations proposées. La description des tâches et leur décomposition éventuelle en sous-tâches définissent un sous-domaine pertinent parmi les domaines représentés afin de réduire les informations accessibles pour chacune d'elles. Ces modèles de domaines ainsi que les connaissances sur les tâches correspondent à une partie des connaissances explicites et partagées de différentes communautés de pratiques – une par discipline. Cet hypermédia est généré dynamiquement à partir des interactions avec l'utilisateur. L'organisation est fondée sur les différents modèles de domaines – disciplines. Les informations sont indexées à partir de ces modèles. La sélection des informations est réalisée par navigation dans un modèle domaine. Il y a bien réutilisation des informations, mais localement. Il n'y a pas ici de modèle utilisateur ni d'adaptation. La sélection des informations pertinentes est fonction de la discipline et de la tâche à réaliser. Nous avons proposé une organisation des connaissances en différentes catégories faiblement couplées - connaissances conceptuelles de domaines provenant des disciplines médicales concernées (neurologie, neurochirurgie, neurophysiologie, ...) et des connaissances contextuelles décrivant des tâches munies de leurs domaines ou sous-domaines pertinents, ainsi qu'une organisation des domaines et sous-domaines.

- Les connaissances sont donc divisées en deux parties : les connaissances de domaine et les connaissances contextuelles. Les connaissances de domaines sont représentées par une base de connaissances intensionnelle. Ces connaissances de domaine nécessitent l'analyse de plusieurs disciplines (scientifiques, médicaux,...) ou plusieurs types d'expertise. En représentation de connaissances, l'importance des systèmes intensionnels a été de nombreuses fois démontrée. Un système est généralement dit intensionnel s'il permet à une même entité d'être représentée de plusieurs façons différentes. Chaque représentation de l'entité possède une même extension mais est caractérisée par des propriétés différentes (intension). Les diverses représentations d'un concept sont reliées par une relation d'égalité extensionnelle, appelée, suivant les auteurs, "coréférence" , "passerelle" , ce qui permet de donner une sémantique unique à chaque représentation et de distinguer explicitement des représentations extensionnellement égales mais intensionnellement différentes.

ATLAS-TUTOR a principalement pour objet d'étudier l'apport de connaissances anatomiques et spatiales – qualitatives - en neuroanatomie pour la conception d'un système d'enseignement assisté par ordinateur. Il s'agit d'aider des radiologistes à décrire et interpréter des images de résonance magnétique du cerveau pour leur formation et ainsi offrir une assistance au diagnostic. Le système MR Tutor développé au Sussex a une connaissance médicale limitée, mais fournit une interface multimédia pour l'élicitation des descriptions d'images par un utilisateur. Le but est donc d'étudier une représentation de connaissances

permettant au système de répondre à des descriptions incorrectes d'images faites par des "étudiants" à différents niveaux d'expertise. Les experts radiologues acquièrent une bonne compétence en intégrant des connaissances médicales et pratiques. Cette intégration est un processus actif qui nécessite : a) d'articuler des connaissances médicales ; b) de les répercuter sur des études de cas ; c) d'établir des connexions entre ces connaissances médicales et les expériences pratiques. Dans notre cas, il s'agit d'étudier les connexions entre des connaissances anatomiques et spatiales sur les structures du cerveau et des lésions, et la description et l'interprétation des images.

Ce chapitre est organisé autour de 5 paragraphes. Le premier décrit la problématique des atlas du cerveau en version papier ou informatisée afin d'en donner les caractéristiques principales, les modes d'utilisation et les besoins. Le second paragraphe décrit plus particulièrement le projet ATLAS pour le traitement des épilepsies focales. Le troisième paragraphe présente le projet ATLAS-TUTOR qui constitue une suite d'ATLAS appliqué à la formation de neuroradiologistes. Les deux derniers sont constitués de la conclusion et de la bibliographie.

2.2 Introduction

La pratique neurologique et neurochirurgicale et un certain nombre de disciplines de recherche relevant des neurosciences conduisent quotidiennement les praticiens et les chercheurs à se référer à des connaissances concernant l'anatomie et le fonctionnement du cerveau. Une façon classique d'organiser ces connaissances consiste à s'appuyer sur le substrat anatomique et sur son mode de représentation le plus naturel, à savoir l'image. Ce principe général est à la base de l'organisation de la plupart des atlas - par exemple : (Schaltenbrand and Wahren 1977; Szikla, Bouvier et al. 1977; Talairach and Tournoux 1988; Ono, Kubik et al. 1991). Ceux-ci contiennent donc des planches anatomiques (constituées à partir de photographies de coupes de cerveaux post-mortem ou d'images anatomiques obtenues in vivo) qui servent de point de départ à une caractérisation textuelle des structures anatomiques (morphologie, relations avec les structures voisines, rôle à l'intérieur d'un système, variabilité, etc).

L'atlas intervient généralement comme support pour l'interprétation de l'anatomie d'une région chez un sujet donné, voire pour la résolution d'un problème physiopathologique ; il peut être utilisé par des experts, des cliniciens, ou même des étudiants. De façon générale, des études en psychologie ont montré que le raisonnement médical nécessite une bonne intégration entre des connaissances générales du domaine considéré et des connaissances en situation (ou connaissances cliniques). Les expériences effectuées par Boshuizen et Schmidt (Boshuizen and Schmidt 1993) ont prouvé que les médecins se servent de plusieurs catégories de connaissances, y compris des connaissances générales (appelées connaissances biomédicales) et des connaissances en situation dérivées des expériences précédentes. Ils ont également démontré que les novices utilisent plus les connaissances biomédicales que les experts. Cependant, dans les cas difficiles où les schémas habituels de raisonnement peuvent échouer, les experts utilisent aussi les connaissances biomédicales. Ces travaux montrent la nécessité d'associer plusieurs catégories de connaissance dans un atlas, à savoir les connaissances biomédicales et les connaissances en situation. D'autre part, ils suggèrent que le système devrait être conçu pour : (1) fournir à l'utilisateur la connaissance la plus

appropriée en fonction de sa tâche et de son niveau d'expertise ; (2) aider à l'émergence des nouvelles connaissances biomédicales à partir des expériences précédentes.

Les connaissances biomédicales du cerveau, définissent les propriétés partagées par la plupart des individus, ou pour des catégories particulières d'individus (droitier, etc.). Elles se composent des quatre catégories suivantes :

- Des connaissances conceptuelles d'une ou plusieurs disciplines du cerveau : elles décrivent et organisent les entités abstraites du cerveau (structures anatomiques, systèmes fonctionnels, neurochimiques, neurophysiologiques, etc.) et les relations entre ces entités (généralisation/spécialisation, spatiales, partie tout, etc.).
- Des illustrations et des textes décrivant des propriétés communes aux cerveaux
- Des données numériques, fonctions de l'espace ou du temps : par exemple les cartes de probabilité 3D représentent la probabilité qu'un point appartienne à une structure anatomique particulière.
- Des modèles de déformation qui permettent de recaler des images de cerveaux différents, des algorithmes de segmentation d'images, etc.

Les connaissances en situation (cas) : elles caractérisent des individus particuliers dont les données ont été enregistrées (par exemple des images, des signaux physiologiques) et interprétées, afin de rendre leur caractère spécifique explicite. Elles peuvent se composer de :

- Données numériques, fonctions de l'espace ou du temps : par exemple des images ou des signaux physiologiques.
- Description de cerveaux dont les attributs décrivent les caractères spécifiques de chacun (par exemple longueur et profondeur d'une sulcature corticale, volume de matière grise dans un gyrus), ou des relations spatiales entre des structures anatomiques.

L'utilisation d'un atlas s'inscrit tout à fait dans ce cadre général et permet la mise en correspondance de ces deux types de connaissances : de façon schématique, ce processus actif prend la forme suivante (cf. Figure 1), qui fait apparaître trois étapes :

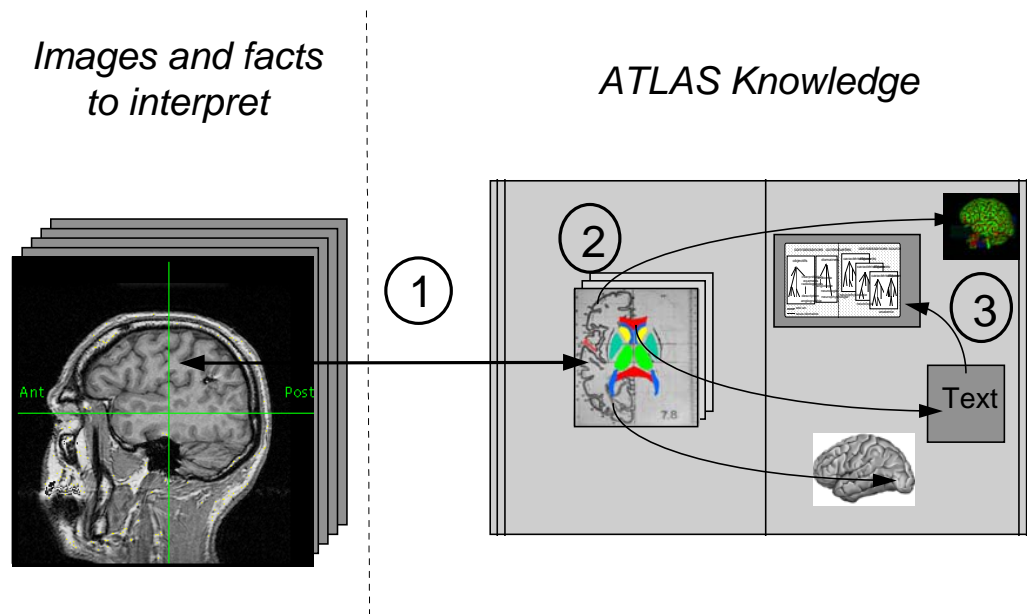


Figure 1: Processus d'utilisation d'un Atlas

1. La mise en correspondance entre l'anatomie du sujet à interpréter et l'anatomie d'un ou plusieurs sujets décrits dans l'atlas au moyen de planches.
2. L'identification des structures chez le sujet à interpréter à partir des délimitations effectuées sur l'atlas et des étiquettes associées.
3. L'accès aux connaissances diverses associées aux structures anatomiques identifiées.

Un atlas numérique du cerveau permet de proposer de nouvelles fonctionnalités :

- La première concerne le mode de représentation des images. L'atlas numérique permet quant à lui d'associer au jeu de planches 3D un outil logiciel de visualisation 3D capable d'offrir simultanément des représentations en coupes et/ou sous forme de surfaces, qui sont toutes les deux nécessaires à la compréhension (cf. Figure. 2).
- La seconde concerne la prise en compte de la variabilité interindividuelle. En raison de la grande variabilité interindividuelle, il est nécessaire de comparer un cerveau à interpréter avec un ou plusieurs cerveaux de référence (cf. Figure. 3). L'atlas numérique rend possible l'utilisation de modèles plus complexes (utilisant des modèles élastiques au lieu d'interpolations linéaires), voire l'utilisation conjointe de différents modèles (plus ou moins locaux donc plus précis). En outre une gestion plus explicite du degré de précision offert devient possible.
- La troisième concerne l'extensibilité. Contrairement au livre traditionnel, un atlas numérique est extensible : cette extensibilité concerne : i) le jeu de planches utilisables (multiplier le jeu de planches disponibles est une forme de réponse au problème de la variabilité interindividuelle) ; ii) les modèles de déformation utilisables ; iii) l'ensemble des informations accessibles.

- La quatrième concerne la gestion cohérente d'informations représentées sous des formes diverses, en particulier sous des formes appréhendables par l'homme (illustrations de type textes ou schémas, images 3D - provenant notamment de patients précédemment analysés, connaissances symboliques) ou par l'ordinateur (connaissances symboliques et/ou images et/ou textes). La notion de cohérence est ici directement liée aux activités des utilisateurs, par exemple un neuroanatomiste n'accède pas aux mêmes informations qu'un neurophysiologiste et elles ne sont généralement pas présentées de la même manière.
- La cinquième concerne l'accès à l'information. La masse d'information disponible est très importante. Il n'est donc pas possible de proposer à chaque utilisateur d'accéder à toutes les informations disponibles. Il s'agit plutôt de limiter les informations accessibles à celles considérées comme pertinentes au regard du contexte, défini à partir du profil de l'utilisateur (expert, clinicien ou étudiant), du domaine médical concerné (neurologie, neuroanatomie, neurochimie, etc.) et de la tâche à réaliser (Waern 1986). Cet accès à l'information pertinente est directement lié à la question de la gestion cohérente de l'information. Il s'agit en effet de filtrer l'accès aux informations et donc de ne conserver que celles qui sont pertinentes. Ceci permet, notamment, de fournir à l'utilisateur parmi différentes descriptions possibles d'un concept relatif au cerveau, celle qui est adaptée au contexte (Chaffin 1981; Roth and Shoben 1983; Barsalou and Medin 1986; Medin and Shoben 1988).

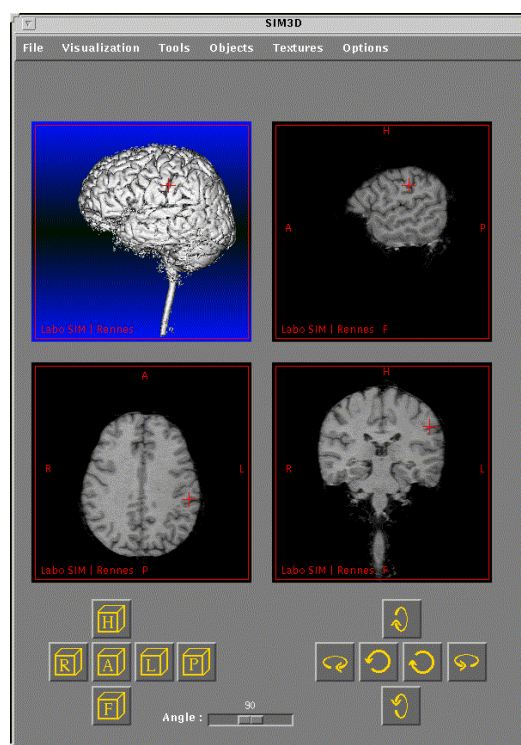


Figure 2 : Exemple d'affichage en 3D.

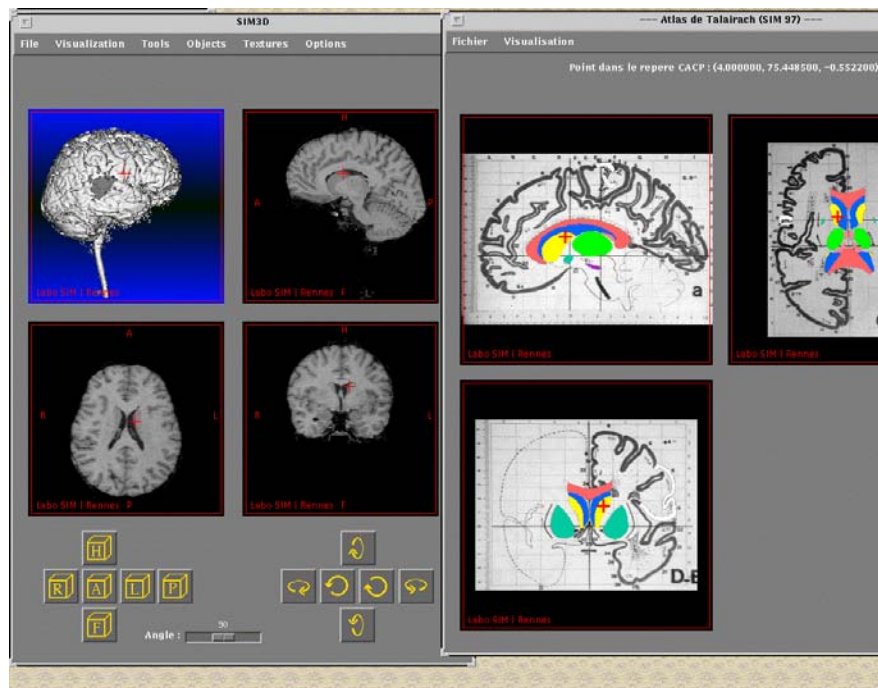


Figure 3: Interprétation et appariement à partir d'un atlas.

Depuis quelques années on assiste au développement d'atlas informatiques : un grand nombre de travaux visent à modéliser la variabilité interindividuelle, par exemple : (Bajcsy and Kovacic 1989; Bookstein 1991; Evans, Collins et al. 1992; Gee, Barillot et al. 1994). Graf von Keyserlingk a numérisé et étiqueté un ensemble de cerveaux post-mortem dans le but d'étudier la variabilité des gyri du cortex (Keyserlingk, Niemann et al. 1985; Keyserlingk, Niemann et al. 1988). Certains auteurs ont tenté d'intégrer ces modèles dans de véritables atlas, notamment : (Greitz, Bohm et al. 1991). Parallèlement, Toga a développé un atlas 3D anatomique et fonctionnel des structures anatomiques cérébrales chez le rat (Toga 1991). A partir d'un prototype appelé "Brain Browser" (Bloom 1990), Bloom jette les bases d'une nouvelle génération d'atlas, capables de représenter l'organisation en systèmes d'un grand nombre d'entités, en mettant l'accent sur la représentation de ces entités sous forme d'un modèle symbolique (Niggemann 1990). Par ailleurs, de nombreux auteurs se sont intéressés à la constitution d'atlas permettant d'aider à l'interprétation de données anatomiques, par exemple (Sokolowska and Newell 1986; Natarajan, Cawley et al. 1991; Zachmann 1991; Robinson, Colchester et al. 1993; Staemmler, Claridge et al. 1993), ou de données fonctionnelles : PET, SPECT et MEG, par exemple (Fox, Perlmutter et al. 1985; Friston, Passingham et al. 1989; Seitz, Bohm et al. 1990; Bohm, Greitz et al. 1991; Evans, Marrett et al. 1991; Lehmann, Hawkes et al. 1991). Avec les mêmes objectifs généraux, le projet américain présenté par Mazziotta - contribution au « Human Brain Project » - met en avant la constitution d'une base de données multicentrique, afin d'évaluer la variabilité morphométrique du cerveau (Mazziotta, Toga et al. 1993). Enfin, certains systèmes ont été conçus autour d'outils de représentation 3D (Höhne, Bomans et al. 1992).

Tous ces travaux permettent soit de définir des modèles plus adéquats de la variabilité interindividuelle, soit de préciser comment les outils de représentation 3D peuvent être utilisés, soit d'illustrer comment les connaissances de l'atlas peuvent être utilisées pour

résoudre des problèmes d'interprétation d'images. Deux aspects importants paraissent cependant ne pas avoir été étudiés :

1. Le premier concerne les connaissances spatiales qui sont présentes dans les images anatomiques, dans les textes relatifs aux relations spatiales entre structures anatomiques et dans leur caractérisation par des neuroanatomistes ou neurochirurgiens. Un modèle de l'espace se doit d'être le plus proche possible de la façon dont il est perçu et décrit par les experts, cliniciens ou étudiants, mais aussi formalisable par des structures mathématiques abstraites afin de pouvoir plus tard servir de base à du raisonnement spatial (Downs and Stea 1973; Hart and Moore 1973; Freksa 1991; Mark and Frank 1991; Randell, Cui et al. 1992; Cohn 1993; Cui, Cohn et al. 1993; Hernandez 1993; Hernandez and Zimmermann 1993; Cohn and Gotts 1994; Cohn, Randell et al. 1994; Hernandez 1994). Cet aspect a été étudié dans le projet Atlas-Tutor pour la détection d'erreurs chez l'apprenant.
2. Le second concerne l'accès aux informations/connaissances (images, illustrations, connaissances symboliques). Pour l'interprétation de l'anatomie, le système ne peut se substituer aux neuroanatomistes, neurologues ou neurochirurgiens. Il s'agit en effet de tâches qui requièrent une réelle expertise :
 - Elles mettent en œuvre des processus d'interprétation d'images qui nécessitent une longue pratique non formalisable. En d'autres termes, un atlas informatique ne peut proposer que des solutions partielles - des résultats intermédiaires, seul le spécialiste a une réelle maîtrise de ce type de processus.
 - Les connaissances sont incomplètes et font appel à de nombreuses disciplines. Les interactions entre ces disciplines pour la résolution du problème sont mal connues. Les stratégies de résolution ne sont pas formalisables.
 - les décisions prises pendant la résolution du problème doivent souvent être remises en cause.

En conséquence, la répartition de l'action et de l'interprétation se fait de la manière suivante : le processus de « résolution de problème » est entièrement contrôlé par l'utilisateur – le médecin. En effet, le processus peut être inconnu de l'utilisateur lui-même : il est habituellement très heuristique et progressif et peut exiger des étapes de retours arrière fréquents. Pour toutes ces raisons, un atlas informatisé ne peut que proposer des mécanismes de recherche et d'accès aux informations/connaissances pertinentes et aux outils de visualisation et de traitement d'images relativement à la tâche ou au domaine de l'utilisateur en garantissant la cohérence entre toutes les sources d'information et de connaissances. Le projet ATLAS a principalement contribué au second aspect tandis que le projet ATLAS-TUTOR a étudié l'apport des atlas informatisés du cerveau à l'environnement MR-TUTOR du projet MEDiate pour l'enseignement assisté par ordinateur. Il s'agit de former des neuroradiologistes au diagnostic de lésions du cerveau en détectant les erreurs de diagnostic et en procédant à une remédiation en utilisant un atlas informatisé et du raisonnement spatial.

2.3 Le projet ATLAS

Ce projet vise à concevoir un système d'information sémantique, encore appelé atlas informatique du cerveau et concerne la recherche fondamentale sur le fonctionnement cérébral, mais, également, les applications médicales, à des degrés multiples, allant depuis l'amélioration de techniques de segmentation d'images médicales numériques (par l'utilisation de modèles de connaissances sur les structures à détecter) jusqu'à la réalisation de systèmes d'aide au diagnostic ou d'aide à la décision thérapeutique. Nous nous sommes principalement intéressés aux activités suivantes :

- L'aide à l'étiquetage des structures corticales (gyri et vaisseaux) dans des images IRM et angiographiques, sous des formes manuelle ou assistée : elle inclut la réalisation d'une base de connaissance anatomique. Elle présente un intérêt indiscutable : i) sur le plan clinique, dans une perspective de planning thérapeutique assisté par ordinateur (neurochirurgie stéréotaxique ou radiochirurgie) et ii) sur le plan de la recherche, pour progresser dans la mise en œuvre d'algorithmes d'interprétation d'images fondés sur des connaissances symboliques.
- L'étude des corrélations morphologiques et fonctionnelles dans la chirurgie de l'épilepsie : cette dernière est plus prospective : elle inclut la définition d'une base de connaissances anatomique et fonctionnelle, susceptible d'aider à confirmer ou à rejeter des hypothèses relatives à la mise en jeu de régions cérébrales pendant le déroulement d'une crise d'épilepsie. Ces deux aspects relèvent principalement des domaines de la neuroanatomie, de la neurophysiologie et de la neurologie. Néanmoins, il pourra s'avérer nécessaire d'inclure dans l'avenir d'autres domaines qui interviendraient pour ces traitements (neurochimie, par exemple).

Dans un tel cadre, le système d'information doit permettre à des utilisateurs issus de domaines différents (neurologie, neurochirurgie, neurophysiologie...), et ayant des objectifs différents d'accéder à des connaissances et des informations relatives au cerveau humain. Ce système ne doit fournir à l'utilisateur que les informations pertinentes en fonction du contexte. Cette notion de contexte est souvent interprétée de manière multiple. Mittal et Paris ont analysé la notion de contexte de communication lors des interactions homme/machine qui comprend notamment les critères suivants : la résolution de problèmes (qui peut être décomposée en tâches réalisées, méthodes pour réaliser une tâche, résultat attendu, domaine de représentation et enfin état du système), les participants impliqués (tient compte de l'expertise, des croyances de l'utilisateur, des buts et plans de celui-ci), le mode d'interaction (nombre de participants, moyen d'interaction utilisé, prise en compte par le système du dialogue passé), le discours (but communicatif et moyens utilisés pour communiquer l'information), et enfin le monde externe (type de texte, environnement...) (Mittal and Paris 1993). Dans notre cadre, nous avons retenu les critères suivants pour le contexte : discipline médicale concernée (par exemple, neurophysiologie, neuroanatomie, neurochimie) et tâche à réaliser. Seuls les concepts pertinents dans une discipline sont décrits dans celle-ci, cela permet d'une part de fournir à l'utilisateur parmi différentes descriptions possibles d'un concept relatif au cerveau, celle qui est adaptée au contexte (Chaffin 1981; Roth and Shoben 1983; Barsalou and Medin 1986; Medin and Shoben 1988), et d'autre part cela permet de restreindre les informations accessibles en fonction de la discipline considérée. La description des tâches et leur décomposition éventuelle en sous-tâches définit un sous-domaine pertinent

parmi les domaines représentés afin de réduire les informations accessibles pour chacune d'elles.

2.3.1 Architecture du système d'information

Dans le projet Atlas, le système contient les connaissances biomédicales suivantes :

- Des connaissances conceptuelles de domaine provenant des disciplines médicales concernées (neurologie, neurochirurgie, neurophysiologie, ...) et des connaissances contextuelles décrivant des tâches munies de leurs domaines ou sous-domaines pertinents, ainsi qu'une organisation des domaines et sous-domaines.
- Des images, des textes, des illustrations, des planches de Talairach munies de structures anatomiques étiquetées.
- Des données numériques, fonctions de l'espace ou du temps : par exemple les cartes de la probabilité 3D représentent la probabilité qu'un point appartient à une structure anatomique particulière et des modèles de déformation qui permettent de recaler des images de cerveaux différents, des algorithmes de segmentation d'images, etc. qui sont accessibles à partir d'outils de traitement et de visualisation d'images.

Les connaissances en situation sont constituées des données brutes ou étiquetées (IRM, angiographies, EEG, MEG ...) recueillies lors d'examens réalisés chez des patients ou des sujets sains. Ces deux catégories de connaissances constituent les sources d'information accessibles à partir du système d'information. Les connaissances conceptuelles vont jouer les rôles suivants :

- Les connaissances de domaines sont des sources d'information à part entière. L'utilisateur peut ainsi consulter le contenu des modèles de domaines.
- Les connaissances contextuelles permettent à chaque utilisateur de définir un sous-espace d'information pertinent relativement à sa tâche.
- Les connaissances de domaines indexent toutes les sources d'informations. Chaque information est associée à un domaine et est indexée par tous les concepts et/ou relations du domaine présents dans son contenu et par son type – texte, images, graphe, etc.
- Chaque domaine ou sous-domaine fournit une organisation explicite fondée sur un graphe orienté sans circuit de l'espace d'information associé. Le mécanisme de navigation a pour objet la construction d'une requête - implicite - en parcourant un domaine afin de filtrer l'espace d'information et de sélectionner une information unique.

L'architecture du système d'information est la suivante (cf. Figure. 4) :

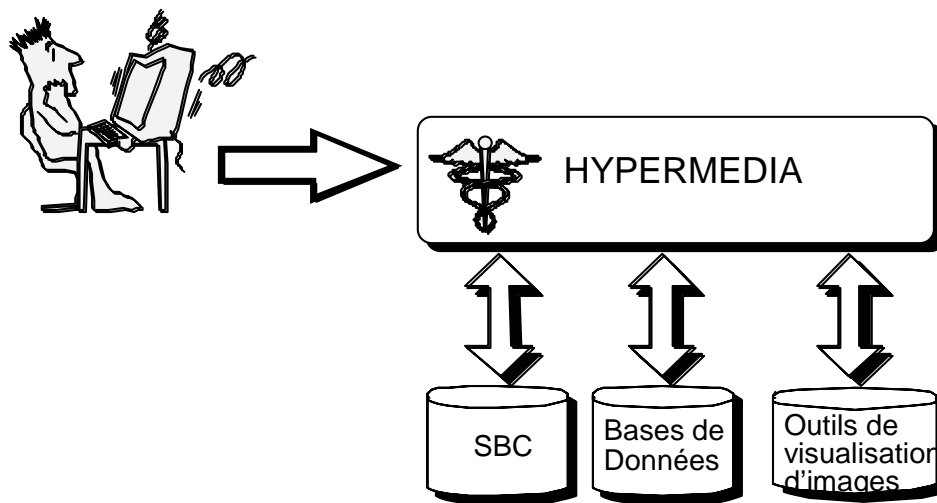


Figure 4 : Organisation générale du système

Le système à base de connaissances contient les connaissances conceptuelles décrites précédemment. Les deux bases de données sont fondées sur un modèle de base de données objet. L'une des bases de données est constituée des images, textes, illustrations et planches de Talairach provenant d'atlas papiers. L'autre base est constituée des données brutes ou étiquetées recueillies lors d'examen et des données numériques, fonctions de l'espace ou du temps. Les outils de traitement et de visualisation d'images 2D permettent de délimiter des zones d'intérêt dans les images des patients en utilisant notamment les modèles de déformation, de les étiqueter en référençant des concepts de la base de connaissances et de permettre leur visualisation bi- ou tridimensionnelle. Le système hypermédia permet aux utilisateurs de rechercher des informations pertinentes dans un contexte donné – un domaine ou un sous-domaine particulier, mais aussi de déterminer ce contexte et d'interagir avec les outils de traitement et de visualisation d'images. Les interfaces et interactions entre l'hypermédia, les bases de données et les outils de traitement d'images ne sont pas détaillées ici. Nous nous focalisons sur les principes de l'hypermédia et ses relations avec les connaissances conceptuelles et les modèles des schémas de bases de données objet.

2.3.1.1 L'hypermédia

L'hypermédia propose l'accès aux informations par un mécanisme de navigation proposant des « boucles » d'interaction de type : choix d'un contexte, sélections successives de critères affinant la recherche par filtrage du sous-espace d'information courant jusqu'à obtention d'une unique information (Montabord, Gibaud et al. 1993; Montabord, Garlatti et al. 1994; Montabord, Gibaud et al. 1995; Montabord 1996). En d'autres termes, la navigation permet à l'utilisateur de construire une requête implicite, constituée d'une conjonction de couples (attribut/valeur) et d'un domaine ou contexte. Cette requête implicite est représentée par le concept de point vue. Cette construction se fait à l'aide de la création dynamique et contextuelle de nœuds et de liens qui composent cet hypermédia.

L'hypermédia peut être vu comme un document virtuel dont l'organisation explicite est fondée sur les modèles de domaines, la sélection est assurée par la navigation construisant le point de vue et l'assemblage défini les différentes catégories de nœuds ainsi que leur contenu.

2.3.1.1.1 L'assemblage ou présentation des différents nœuds

Deux types de nœuds existent : des nœuds d'information contenant l'information à visualiser ainsi que des liens et des nœuds d'organisation. Les nœuds d'organisation ont pour but d'aider l'utilisateur à sélectionner une information unique parmi un ensemble présélectionné.

2.3.1.1.1.1 Les nœuds d'information

Les nœuds d'information proposent des interactions et l'affichage d'un type d'information particulier. Les différents types d'informations sont les suivants : graphes, descriptions, textes, illustrations, images, planche de Talairach (cf. Figures 5, 6, 7, 8 et 9). Un nœud d'information se compose d'un contenu – appartenant à un type d'information - et de liens. Il y a deux catégories de liens : des liens contextuels « Domaine » et des liens caractérisant l'information à rechercher (Objet, Attribut, Présentation, Divers). Ces liens permettent respectivement l'accès aux modèles d'organisation des domaines et des tâches et aux classes et relations du modèle de domaine courant - et donc aux différents types d'information et à leur présentation.

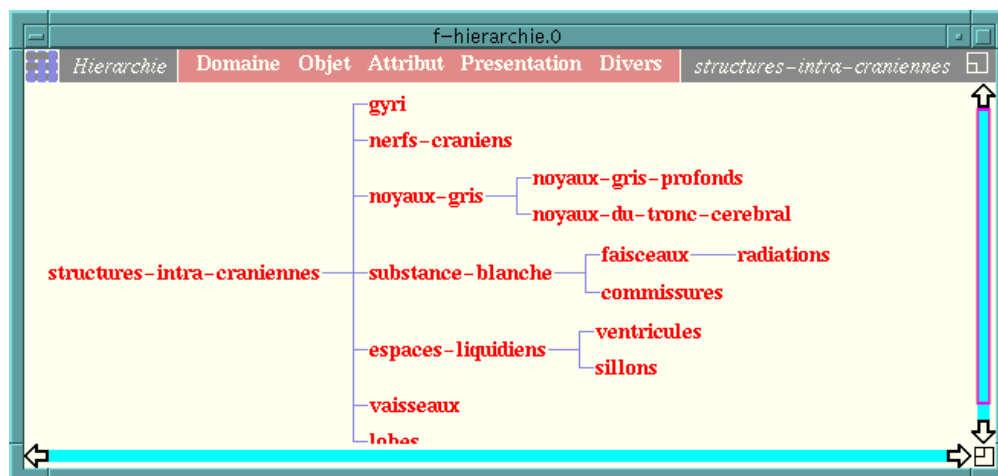


Figure 5 : Type hiérarchie

L'exemple ci-dessus affiche le graphe d'une relation qui est ici le graphe d'héritage entre classe du domaine « neuroanatomie ». Chacun des nœuds de ce graphe est un lien vers la classe correspondante du domaine courant.

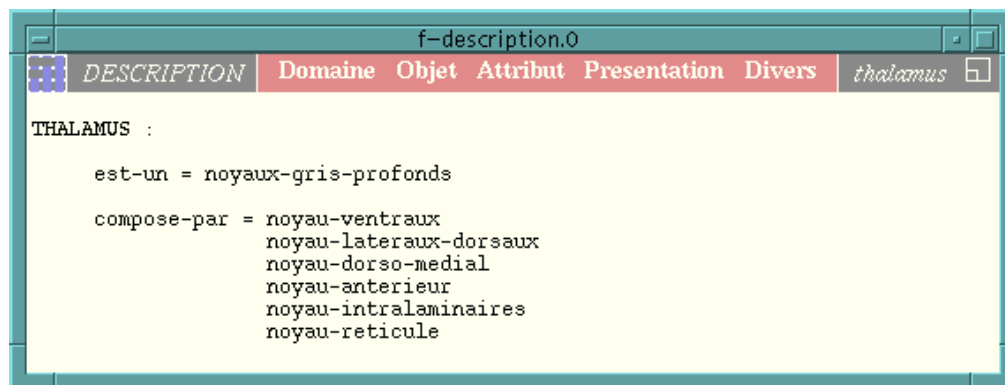


Figure 6 : Type description

L'exemple ci-dessus affiche la définition d'une classe du domaine « neuroanatomie ».

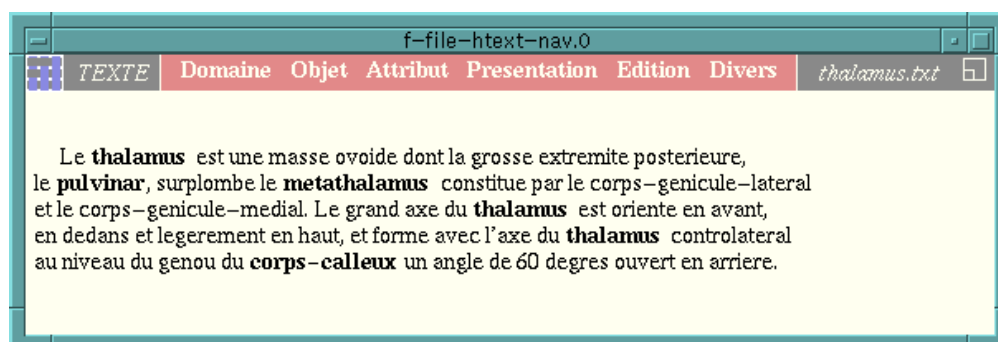


Figure 7 : Type texte

Le type texte affiche un texte provenant d'un atlas papier ou numérisé. Les concepts du modèle de domaine courant présents dans le contenu sont des liens – en gras.

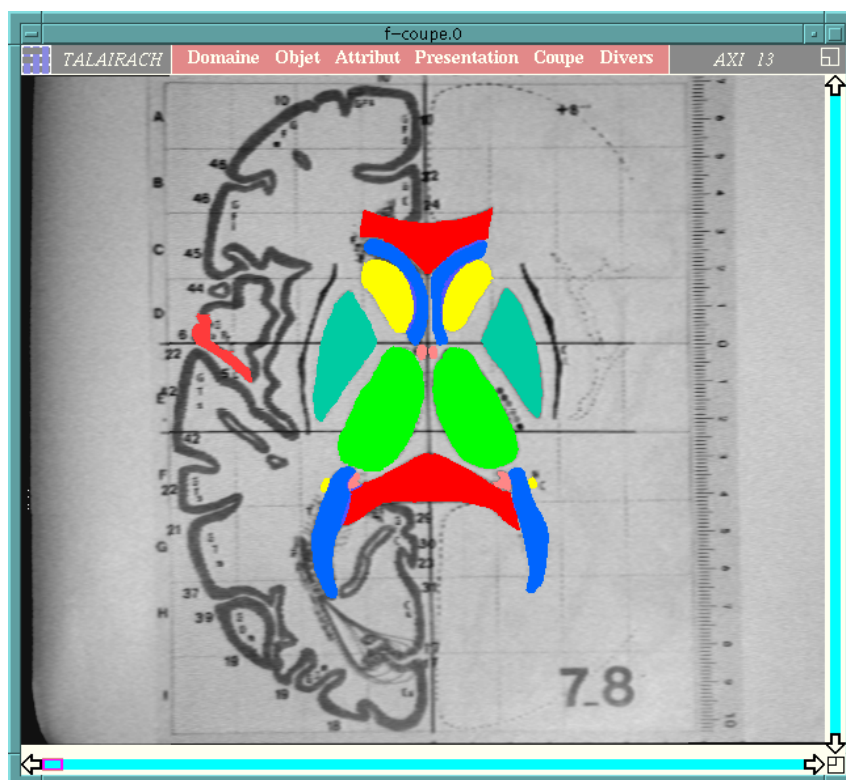


Figure 8 : Type Planche de Talairach

Ce type affiche une coupe d'atlas. Une coupe est munie de structures anatomiques étiquetées et délimitées qui sont des aussi des liens.

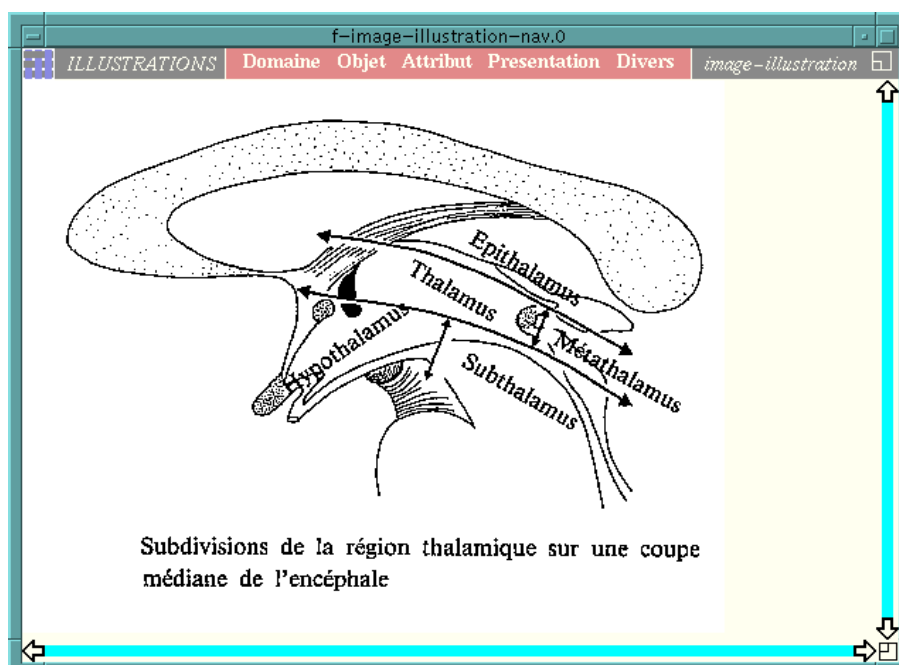


Figure 9 : Type illustration

Le type illustration affiche un dessin provenant d'un atlas papier ou numérisé.

2.3.1.1.2 Les nœuds d'organisation

Les nœuds d'organisation sont uniquement composés de liens permettant d'affiner la ou les sélections précédentes jusqu'à l'obtention d'une information unique. Les descripteurs présents dans ces nœuds sont le résultat de l'analyse de l'espace d'information courant et de ces critères de sélection discriminants (cf. Figure 10). Ces critères sont extraits des modèles de domaines et/ou des schémas de bases de données.

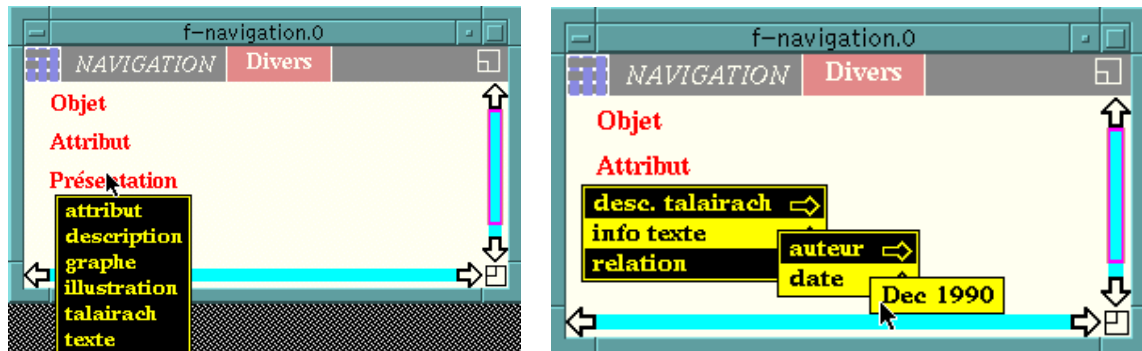


Figure 10 : Nœud d'organisation

2.3.1.1.2 Le mécanisme de navigation

Le mécanisme de navigation permet à l'utilisateur de sélectionner une unique information en construisant une requête implicite à l'aide d'un point de vue. Ce dernier représente l'état courant de la requête. Le point de vue définit le contexte courant et désigne un ensemble d'informations accessibles à partir d'une conjonction de couples (attribut/valeur) qui le compose. Il est représenté par un objet dont les attributs sont appelés critères. Un point de vue est constitué par deux ensembles de critères : les critères contextuels d'une part et les critères de domaines et de type d'informations d'autre part. En d'autres termes, il offre la possibilité de déterminer toutes les propriétés qui indexent les informations. Un point de vue est associé à chaque nœud d'information ou de navigation. A chaque sélection de lien, un nouveau point de vue et un nouveau nœud est créé, sauf quand l'espace d'information est réduit à une information unique. Les attributs du point de vue sont soit de nouveaux critères contextuels soit des critères et des valeurs sélectionnées aux étapes précédentes de la boucle de navigation auxquels s'ajoutent le couple (critère/valeur) correspondant au lien choisi. La liste des points de vue ayant amené à une étape de la navigation est associée à un nœud d'organisation afin de pouvoir revenir en arrière dans le processus de sélection.

Le mécanisme de navigation est constitué d'une suite de cycles. Un cycle se déroule en quatre étapes (cf. Figure 9) :

- Choix d'un lien par l'utilisateur.
- Création d'un nouveau point de vue. Ce point de vue est associé au nœud qui sera créé lors de l'affichage de l'information sélectionnée ou du nouveau nœud d'organisation.
- Sélection des informations accessibles à partir du nouveau point de vue (filtrage). Les bases données et les connaissances de domaine sont interrogées séparément, avec le

point de vue comme filtre. Les informations accessibles dans chaque source d'information sont ensuite fusionnées.

- Création d'un nœud. Si une information unique est accessible alors un nœud d'information est créé, sinon un nœud d'organisation est créé afin de permettre à l'utilisateur de préciser son choix. Le point de vue courant est associé au nœud. Lors de la création d'un nœud d'information, seuls les critères contextuels (définis plus loin) du point de vue sont associés au nœud créé. En effet, les critères de domaines qui ont permis de sélectionner l'information ne sont plus nécessaires. Les liens proposés dans un nœud d'organisation sont ceux qui fournissent les couples (attribut/valeur) discriminant les informations accessibles à partir du point de vue courant.

Ce mécanisme de navigation est fondé sur des connaissances de domaines multiples (différentes disciplines intervenant pour la préparation d'actes médicaux pour la chirurgie de l'épilepsie) et des connaissances contextuelles. Nous allons maintenant expliciter l'organisation de ces connaissances et les mécanismes d'inférence associés.

2.3.2 Architecture de la base de connaissances.

La base de connaissances doit prendre en compte l'évolution des disciplines médicales, tant par leur contenu (ensemble de connaissances pertinentes dans une discipline), que par leur nombre (adjonction de nouvelles disciplines). Dans un tel cadre, des sources multiples de connaissances et/ou des représentations multiples procurent un excellent moyen pour structurer la base de connaissances (Simmons and Davis 1993). Résoudre un problème nécessite souvent différents types de connaissances (Simmons and Davis 1993) telles que les connaissances concernant les tâches, les mécanismes d'inférences et les connaissances de domaine. Chacun de ces types peut lui-même être partitionné. L'utilisation de connaissances sous une forme modulaire permet une extensibilité plus aisée du système pour traiter de nouvelles tâches ou pour les traiter de façon plus précise. Partitionner les connaissances de domaine est également important car il est souvent difficile de modéliser tous les aspects du problème considéré dans une organisation unique. Il est bien plus intéressant de diviser les connaissances de domaine en plusieurs parties quasi-indépendantes. En effet, lorsque les sources de connaissances sont peu dépendantes les unes des autres, les connaissances du système peuvent être augmentées par adjonction de nouvelles connaissances, ou en affinant les connaissances déjà présentes sans modifier le reste du système. Séparer les différents types de connaissances procure des avantages en acquisition de connaissances, extension du système, réutilisation et compréhension.

Nous proposons donc une organisation de la base de connaissances en différents types de connaissances faiblement couplées - connaissances conceptuelles de domaines provenant des disciplines médicales concernées (neurologie, neurochirurgie, neurophysiologie, ...) et des connaissances contextuelles décrivant des tâches munies de leurs domaines ou sous-domaines pertinents, ainsi qu'une organisation des domaines et sous-domaines. La base de connaissances est donc divisée en deux parties : les connaissances de domaine et les connaissances contextuelles (cf. Figure 11). Les premières sont organisées en plusieurs perspectives quasi disjointes (une par discipline médicale concernée). Les secondes contiennent un modèle hiérarchique de tâches ainsi que des connaissances sur les disciplines médicales. Ces dernières ont pour objet de délimiter le sous-ensemble de connaissances de

domaines pertinentes pour une discipline médicale, celles-ci s'organisant en disciplines et sous-disciplines.

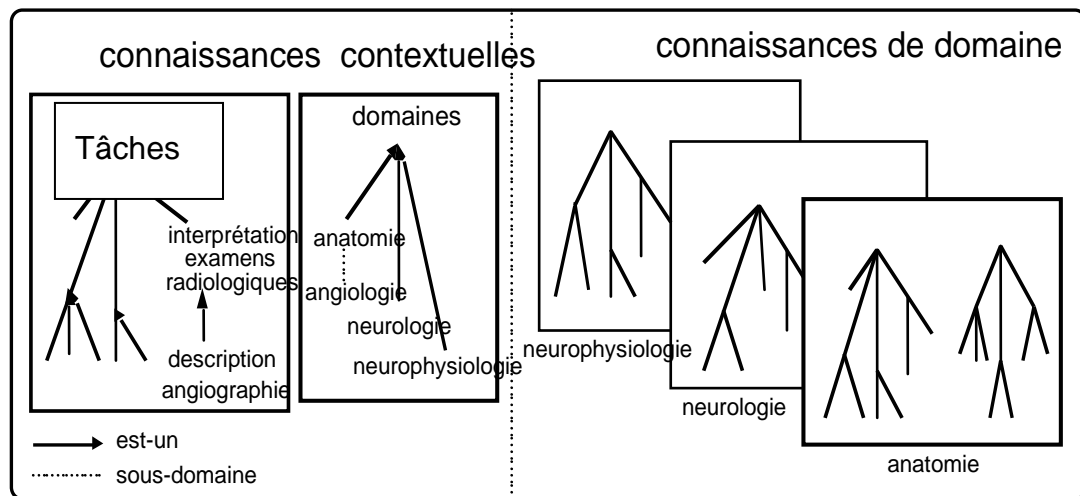


Figure 11 : Organisation de la base de connaissances dans ATLAS

Nous allons maintenant détailler l'organisation des connaissances de domaine.

2.3.2.1 Les connaissances de domaine

Elles sont représentées par une base de connaissances intensionnelle. Ce type de connaissances nécessite l'analyse de plusieurs disciplines (scientifiques, médicaux,...). En représentation de connaissances, l'importance des systèmes intensionnels a été de nombreuses fois démontrée (Woods 1975; Ferber 1988; Ferber and Volle 1988; Marino 1989; Marino, Rechenmann et al. 1990). Un système est appelé intensionnel s'il permet à une même entité d'être représentée de plusieurs façons différentes, c'est-à-dire que chaque représentation de l'entité possède une même extension mais est caractérisée par des intensions différentes (intension : ensemble de propriétés). Les diverses représentations d'un concept sont reliées par une relation d'égalité extensionnelle, appelée, suivant les auteurs, "coréférence" (Brachman 1985; Ferber 1988; Ferber and Volle 1988), "passerelle" (Marino 1989; Marino, Rechenmann et al. 1990), ce qui permet de donner une sémantique unique à chaque représentation et de distinguer explicitement des représentations extensionnellement égales mais intensionnellement différentes.

Une base de connaissances, ici fondée sur le paradigme objets, est constituée d'un ensemble d'objets structurés par un ensemble de relations. La relation de généralisation / spécialisation organise principalement les objets entre eux. Le mécanisme d'héritage ou de partage d'attributs, défini au dessus de cette relation, est ici au centre de nos préoccupations. En effet, il s'agit d'autoriser les conflits d'héritage multiple. Il existe deux types de conflits d'héritage : les conflits de valeurs et les conflits sémantiques ou conflits de nom. Le premier se produit lorsque les attributs des classes en conflit dénotent la même entité, c'est-à-dire que ce sont les mêmes attributs. Le second se produit quand les attributs des classes en conflit dénotent une entité différente, c'est-à-dire que ce sont des attributs différents ayant le même

nom. Les différents travaux sur les systèmes à héritage multiple montrent que chaque type de conflit peut surgir dans des représentations (Bobrow and Winograd 1977; Bobrow and Stefik 1981; Borning 1986; Dugerdil 1986; Carré and Comyn 1988; Chouraqui and Dugerdil 1988; Knudsen 1988; Carré and Comyn 1989; Ducournau and Habib 1989; Carré, Dekker et al. 1990; Carré and Geib 1990; Ducournau, Habib et al. 1995). Il est bien évident que si l'on renomme l'un des attributs, le conflit disparaît. Mais, il est important de conserver dans une représentation les noms utilisés par les domaines d'expertise concernés. Par contre la plupart des études ne concernent bien souvent qu'un seul type de conflit et sont donc dans certains cas restrictifs.

Nous proposons une méthodologie de conception de base de connaissances intensionnelle composée de plusieurs perspectives quasi-indépendantes appelées domaines. Les objets de différents domaines qui représentent une même entité sont reliés par une relation d'égalité extensionnelle. Ces domaines nous permettent de traiter la résolution des conflits en deux étapes, la première résout les conflits sémantiques ou inter-domaines et la seconde les conflits de valeur ou conflits intra-domaine. Cette méthodologie a été appliquée pour un système d'aide à la décision sur le cerveau humain afin d'assister des neurochirurgiens pour le traitement de pathologies (Garlatti, Kanellos et al. 1994), faisant intervenir des domaines d'expertise tels que neuroanatomie, neurologie, neurophysiologie, etc.

Dans un premier temps nous introduisons brièvement un cadre général pour l'héritage et les différents types de conflits. Puis les deux principales stratégies de résolution de conflits sont analysées et discutées. Différentes approches de la notion de perspective sont présentées dans le cadre des représentations à objets. Cette notion forme la pierre angulaire de l'architecture de la base de connaissances. Le mécanisme d'héritage prenant en compte les perspectives et les différentes natures de conflits associées est explicité, ainsi que les propriétés de l'algorithme correspondant.

2.3.2.1.1 Les conflits d'héritage multiple

2.3.2.1.1.3 Le cadre général

Une représentation objets est composée d'un ensemble d'objets structurés par un ensemble de relations. Parmi toutes ces relations, l'une d'entre elles est ici au centre de nos préoccupations : la relation de généralisation / spécialisation. La hiérarchie induite par cette relation est appelée graphe d'héritage. Un objet est défini par un ensemble de propriétés. Une propriété est une paire (Attribut, valeur). Les attributs dont les valeurs sont des noms d'objets sont des relations. Seuls les attributs spécifiques à un objet sont déclarés dans ce dernier - les autres étant accessibles par le mécanisme d'héritage ou de partage d'attributs utilisant le graphe d'héritage. Les objets et les attributs sont désignés par un nom. Un attribut peut avoir une ou plusieurs caractéristiques comme un domaine, une cardinalité et une valeur. Dans la suite du papier sauf exception nous ne parlerons que de valeurs quelle que soit la caractéristique afin de simplifier la présentation. Pour définir un cadre général à l'héritage multiple, les notions de hiérarchie, de conflits sémantiques et de valeurs, de masquage et d'ensemble de conflits sont introduites. Puis les deux principales stratégies de résolution de conflits sont présentées : linéarisation et désignation explicite. Enfin, ces deux stratégies sont discutées.

Un graphe d'héritage, noté $G = (X, H)$, est un graphe orienté sans circuit où X est l'ensemble des objets et H l'ensemble des liens. X se compose de deux ensembles disjoints : X' l'ensemble des objets abstraits ou classes et X'' l'ensemble des objets individuels ou instances. H se compose aussi de deux ensembles disjoints : H' l'ensemble des liens entre classes appelé "is-a" et H'' l'ensemble des liens entre classes et instances, appelés "member-of". L'ordre partiel induit par le graphe d'héritage est noté \leq_H .

Définition 1 : $\forall x \in X, \forall y \in X'$, tel que $x \leq_H y$, y est appelé l'ancêtre de x et x le descendant de y .

Une classe x possède une extension propre notée $Ext_p(x)$ qui se compose de toutes les instances qui sont des descendants de x et une extension (au sens large) $Ext(x)$ qui inclut tous ses descendants - classes et instances.

Principe 1 : dans le graphe d'héritage $G = (X, H)$, $x \in X, y \in X'$, tels que $x \leq_H y$, $Ext(x) \subset Ext(y)$ et $x \in X''$ et $y \in X'$ tels que $x \leq_H y$, $x \in Ext_p(y)$.

Ce principe est toujours applicable dans les représentations objets. Une classe donnée ne peut hériter que d'elle-même ou de ses ancêtres. La définition 2 et le principe 2 explicite cette propriété.

Définition 2 : pour un objet donné $\alpha \in X$, son graphe d'héritage (appelé sa hiérarchie) est $G(\alpha) = (X_\alpha, H_\alpha)$ où $X_\alpha = \{y \in X, \alpha \leq_H y\}$, i.e. l'ensemble de ses ancêtres (incluant α lui-même).

Principe 2 : pour un couple donné (x, A) où x est un objet et A un attribut, le mécanisme d'héritage détermine la valeur de A pour x dans $G(x)$.

On fait ici l'hypothèse que le mécanisme d'héritage n'est pas répété et qu'il ne fournit qu'une seule valeur pour A dans $G(x)$ et qu'il satisfait au principe 3 suivant :

Principe 3 : le mécanisme d'héritage est uniforme, i.e. il s'applique de la même manière quel que soit le type d'attribut.

L'un des problèmes majeurs de l'héritage multiple est la résolution des conflits. Dès que deux ancêtres z et y (dans $G(x)$) de l'objet x possèdent un attribut nommé A , un conflit peut exister. L'existence de ce conflit dépend des valeurs associées à z et y et la nature du conflit de l'attribut lui-même. La plupart des mécanismes d'héritage appliquent un principe appelé spécificité ou masquage, c'est-à-dire qu'uniquement les valeurs d'attributs des objets les plus spécifiques sont retenues et peuvent être en conflit. Dans la figure 12, pour la paire (x, A) , seules les classes z et y sont retenues parce qu'elles sont respectivement plus spécifiques que z_1 et y_1 . Ces classes z et y ne sont en conflit que si les valeurs associées à l'attribut A sont incompatibles (domaines et/ou cardinalités sans intersection, valeurs différentes).

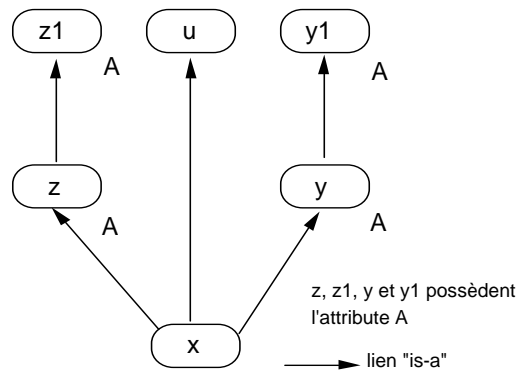


Figure 12 : Conflits d'héritage multiple

Les objets z et y peuvent posséder un même attribut nommé A ou deux attributs différents avec un même nom A . Avec un même attribut, le conflit est appelé conflit de valeur (Ducournau, Habib et al. 1995). Lorsque les attributs sont différents, le conflit est appelé conflit sémantique parce qu'ils dénotent des entités différentes. Ducournau et al (Ducournau, Habib et al. 1995) les appellent conflits de nom et Carré (Carré and Comyn 1989; Carré, Dekker et al. 1990; Carré and Geib 1990) attributs homonymes. Knudsen (Knudsen 1988) a introduit deux classes de conflits nommés "intended name conflict" et "casual name collision" qui recouvrent partiellement chaque type de conflits - valeur et sémantique.

Ducournau et al (Ducournau and Habib 1989) ont proposé un cadre général pour l'héritage multiple. Celui-ci peut être appliqué à presque tous les systèmes et est fondé sur les deux notions suivantes : le masquage et l'ensemble de conflit.

Définition 3 : soient x un objet, A un attribut et y, z deux ancêtres de x dans $G(x)$ qui possèdent l'attribut A :

- z masque y de x pour l'attribut A si $z \leq_H y$ et z est le premier objet possédant (déclarant A avec une caractéristique) A dans le chemin $[x, \dots, z \dots y]$;
- z et y sont en conflit pour A s'ils ne sont pas masqués de x pour A ;
- l'ensemble de conflit $CS(x, A)$ contient tous les objets de $G(x)$ qui possèdent A et ne sont pas masqués de x (i.e. ils sont incomparables par \leq_H).

Cette définition amène tout naturellement le principe 4.

Principe 4 : pour une paire donnée (x, A) , le mécanisme d'héritage obtient une valeur d'un objet appartenant à $CS(x, A)$.

Les stratégies de résolution de conflits sont présentées à partir de ce cadre général et satisfont les principes énoncés ci-dessus.

2.3.2.1.2 Les différentes stratégies

Parmi tous les systèmes à héritage multiple autorisant les conflits, on peut distinguer deux approches majeures : la première applique une stratégie de linéarisation et la seconde une

stratégie de désignation explicite. Nous allons maintenant introduire et discuter ces deux approches.

2.3.2.1.2.4 La linéarisation

Dans une stratégie de linéarisation, l'ordre partiel induit par le graphe d'héritage est transformé en un ordre total ou linéarisation. Pour une paire donnée (x, A) , la valeur de l'attribut A est obtenue dans le premier objet possédant A dans la linéarisation de $G(x)$. Afin de s'assurer qu'un tel mécanisme satisfait le principe de masquage, l'ordre total doit respecter l'ordre partiel de $G(x)$. Seules les extensions linéaires respectent ce principe. Ainsi, le mécanisme d'héritage choisit toujours un objet appartenant à l'ensemble de conflit et satisfait donc le masquage. En fait, une extension linéaire évite le calcul de l'ensemble de conflit. Cette stratégie a été appliquée dans les systèmes suivants : Y3 (Ducournau and Habib 1989) et CLOS (DeMichiel and Gabriel 1987). Mais, il existe plusieurs extensions linéaires qui conservent le même ordre partiel. Des contraintes doivent être ajoutées afin de n'obtenir qu'une seule extension linéaire pour un ordre partiel donné (Ducournau and Habib 1989).

L'avantage principal de cette stratégie est qu'elle permet un calcul automatique de la valeur de l'attribut quelque soit les modifications dans la hiérarchie considérée. En effet, toute modification des déclarations d'attribut est prise en compte lors du parcours de l'ordre total. En cas de modification de la hiérarchie, la linéarisation est recalculée. Cette stratégie présente deux inconvénients : son incapacité à traiter les conflits sémantiques et son manque de flexibilité. En effet, elle n'offre aucun moyen de sélectionner un attribut parmi les autres lors de conflits sémantiques. Pour deux paires $(x, A1)$ et $(x, A2)$, si deux classes appartiennent aux deux ensembles de conflit correspondants, la classe choisie sera la même. En représentation de connaissances, cet ordre immuable entre les classes en conflit n'est pas satisfaisant pour tous les attributs. Ducournau et al (Ducournau, Habib et al. 1995) montrent qu'une telle stratégie est appropriée pour l'héritage de méthodes et de démons et procure un mécanisme d'héritage de base et "par défaut". En effet, il respecte toujours le principe de masquage, c'est-à-dire l'ordre partiel du graphe. Cette dernière n'est pas suffisante pour traiter les conflits sémantiques et n'est pas assez flexible. Par contre, la stratégie de désignation explicite va permettre de traiter les conflits sémantiques.

2.3.2.1.2.5 La désignation explicite

La stratégie de désignation explicite, souvent appelée point de vue, n'est généralement utilisée que pour résoudre des conflits sémantiques. La notion de point de vue a pour objet de choisir l'une des interprétations de l'attribut considéré. Pour une paire (x, A) donnée, x ne peut hériter que d'une classe de $G(x)$. Le point de vue détermine donc un sous-graphe de $G(x)$ ne possédant qu'une seule interprétation de A afin de résoudre le conflit et de fournir une valeur pour A . Nous appellerons ce dernier : sous-graphe associé au point de vue. De la définition du point de vue dépend le sous-graphe associé. Nous allons montrer à travers deux exemples, comment le sous-graphe associé au point de vue est déterminé et quelles en sont les conséquences sur le calcul de la valeur d'un attribut.

sémantique, l'entité qu'il dénote ou son interprétation). A partir de ce principe, il est évident qu'en cas de conflit sémantique, il est nécessaire d'hériter d'autant d'attributs qu'il y a d'interprétations. Par contre en cas de conflit de valeur, on ne doit hériter que d'un seul attribut puisque les diverses déclarations en conflit dénotent la même entité. Dans C++ (Stroustrup 1986), Eiffel (Meyer 1988) et Knudsen (Knudsen 1988), les attributs sont dupliqués pour résoudre les conflits de valeurs et violent donc ce principe. De plus, dans Knudsen (Knudsen 1988), le nombre des attributs hérités est défini par l'utilisateur et dépend de la nature du conflit et du sens du lien "is-a" (union ou intersection). Comme nous l'avons vu précédemment, les différents types de conflits de Knudsen recouvrent ceux présentés ici. Dans certains cas, le principe énoncé n'est donc pas satisfait. Comme nous allons voir, la stratégie de désignation explicite est très proche des notions de perspectives étudiées dans la suite. Contrairement à la stratégie de linéarisation, cette dernière ne permet pas de traiter automatiquement toutes les modifications de la hiérarchie. Ce point particulier sera analysé dans la suite. La notion de point de vue, utilisée ici comme stratégie de résolution de conflit est très proche de celle de perspectives - voire même synonyme.

2.3.2.1.3 Les perspectives

En intelligence artificielle, on trouve trois interprétations principales de la notion de perspective : i) perspectives en tant qu'hypothèses - systèmes Omega (Attardi and Simi 1985), ART (Clayton 1985) et CYC (Guha 1990; Guha and Lenat 1990; Guha 1991), ii) vues spatiales (Minski 1975; Davis 1987) et iii) vues partielles d'une entité. Dans cette dernière - qui seule nous intéresse ici -, une perspective permet d'accéder aux propriétés pertinentes d'une entité relativement à un contexte donné - pris ici au sens large. Dans les représentations objets, la notion de perspective est corrélée avec la relation de généralisation / spécialisation. En effet, pour la représentation d'une entité - un objet -, à une perspective donnée correspond un sous-graphe de cette relation. Ce sous-graphe détermine un espace d'attributs qui sont les attributs pertinents de l'entité considérée dans un certain contexte. La détermination du sous-graphe dépend de la définition de la notion de perspective.

Nous allons analyser et discuter les différentes approches de cette notion de perspective. Elle a été abordée de trois manières différentes : premièrement, sous la forme d'un agrégat de propriétés dans une instance, deuxièmement sous la forme d'un objet - classe ou instance - et troisièmement sous la forme d'un ensemble d'objets.

Selon la première approche, les instances peuvent déclarer plusieurs perspectives qui sont des agrégats de propriétés indépendantes les unes des autres. Une perspective appartient à une seule classe et possède son propre espace d'attributs qui est séparé des autres. Généralement, les attributs ayant un même nom mais avec des sens différents peuvent coexister dans des perspectives différentes. L'héritage multiple n'est pas autorisé. Le sous-graphe associé est constitué de la hiérarchie de la perspective considérée. Dans KRL (Bobrow, Kaplan et al. 1977), OWL II (Martin 1979), PIE (Goldstein and Bobrow 1980) et SYSTALK (Wolinski and Perrot 1991), une instance peut être définie avec plusieurs perspectives tandis que dans LOOPS (Bobrow and Stefik 1981) une instance est composée de plusieurs objets, un par perspective. La notion de perspective peut être vue ici comme une alternative à l'héritage multiple. En effet, si l'on ne sépare pas les perspectives dans une instance alors il y a héritage multiple et des conflits sémantiques peuvent surgir.

Dans la seconde approche, les classes ou instances peuvent être des perspectives - MERING III (Ferber 1988; Ferber and Volle 1988). Plusieurs objets peuvent dénoter la même entité et donc avoir une même extension avec des intensions différentes. Chacun de ces objets est une perspective sur cette entité. Ces objets sont reliés par une relation de coréférence, au lieu d'être regroupés dans un seul objet comme précédemment. Les conflits d'héritage ne sont pas autorisés. Le sous-graphe associé à la perspective est défini par la hiérarchie de l'objet considéré. De plus, cette relation de coréférence permet de définir des contraintes entre les propriétés respectives de ces objets.

Dans la troisième approche, - TROPES (Marino 1991; Marino 1993) - une base de connaissances est constituée d'un ensemble de concepts disjoints. Un concept est représenté par plusieurs perspectives qui représentent chacune un domaine d'expertise. Chaque perspective, séparée des autres, est constituée d'une hiérarchie de classes (un arbre). Certaines classes, appartenant à des perspectives différentes, ont des extensions liées par une relation d'inclusion. Ces classes sont reliées par un pont. Lorsque les classes dénotent la même entité, le pont est bidirectionnel et équivaut donc à une relation de coréférence. Des contraintes entre attributs sont également associées à un pont. Chaque perspective - un arbre - possède une racine qui est reliée par un pont aux autres perspectives puisqu'elles représentent une même entité. Chaque classe n'appartient qu'à une seule perspective, tandis qu'une instance appartiendra à une classe par perspective. Dans une perspective, le sous-graphe associé à une classe est composé de sa hiérarchie et celui d'une instance est composé d'une partie de sa hiérarchie, c'est-à-dire celle appartenant à la perspective désignée. L'héritage multiple est autorisé uniquement pour les instances, mais sans conflit.

L'approche de Ferber et Volle - 2ème approche - est très intéressante parce qu'elle fournit un moyen pour distinguer l'intension de l'extension d'une entité dans la représentation et donc de pouvoir exprimer explicitement que plusieurs objets intensionnellement différents dénotent la même entité. Dans les deux premières approches, la notion de perspective est définie localement ; c'est-à-dire comme un ensemble de propriétés, tandis que dans TROPES, la notion de perspective est plus générale et peut amener à une méthodologie de conception d'une base de connaissances. En effet, tous les objets et propriétés décrivant une entité et pertinents dans un contexte donné sont regroupés dans une seule hiérarchie et forment ainsi une structure conceptuelle. Chaque structure conceptuelle est en fait une source de connaissance appelée connaissances de domaine (Simmons and Davis 1993). Dans notre cadre, le principal inconvénient de TROPES et de MERING III est qu'ils n'autorisent pas les conflits d'héritage multiple.

2.3.2.2 Une base de connaissances intensionnelle

Comme nous l'avons vu précédemment, il est important dans une représentation à objets d'offrir une base de connaissances intensionnelle et la notion de perspectives. Dans le cadre du projet ATLAS (Garlatti, Kanellos et al. 1994), il est nécessaire de traiter l'héritage multiple avec des conflits de valeur et des conflits sémantiques. Le principal problème est donc l'articulation entre la notion de perspective et celle de conflit, ainsi que sa nature.

De nombreux problèmes complexes impliquent l'acquisition de connaissances provenant de plusieurs domaines différents - disciplines scientifiques par exemple ou médicales - décrivant souvent chacun les mêmes entités, mais de manière intensionnellement différentes. Dans un tel cadre, chaque domaine peut être analysé séparément et indépendamment afin

d'obtenir une structure conceptuelle par domaine. Ensuite, il s'agit de d'identifier les entités communes à différents domaines. Nous nous proposons de regrouper toutes les représentations d'un domaine dans une perspective, sous la forme d'un sous-graphe d'héritage quasi-indépendant des autres. Toutes les représentations d'une même entité présente dans différents domaines sont reliées entre elles par une relation de coréférence afin d'explicitier cette différence entre intension et extension. Seules les classes seront reliées par une telle relation ; les instances pourront appartenir à plusieurs domaines. Les perspectives seront ainsi faiblement couplées.

Traiter les deux types de conflits implique de les distinguer et de choisir une stratégie adéquate pour chacun. Pour les conflits sémantiques, seule une stratégie de désignation explicite est envisageable. Afin d'avoir une stratégie flexible et adaptable en fonction des attributs considérés, la stratégie de désignation explicite est également utilisée pour les conflits de valeurs. Mais il s'agit maintenant de voir comment ces deux types de conflits interagissent avec les perspectives qui sont organisées par la relation de généralisation / spécialisation supportant l'héritage.

Dans un premier temps, nous allons définir la notion de domaine, en signaler les conséquences sur la nature des conflits et présenter l'architecture de la base de connaissances. Puis le mécanisme d'héritage correspondant est explicité. Enfin, les caractéristiques principales des algorithmes de résolution de conflits sont données.

2.3.2.2.1 Nature des conflits

Avant de présenter la stratégie de résolution des conflits, nous allons d'abord analyser les différents types de conflits existants : les conflits inter-domaines ou conflits sémantiques et les conflits intra-domaines ou conflits sur des valeurs. D'un domaine à un autre, les propriétés qui définissent les concepts sont différentes mais possèdent parfois un même nom. Le mécanisme d'héritage utilise ce nom comme référence pour rechercher la valeur associée à ce nom dans la hiérarchie de l'objet considéré. Comme nous allons le voir, à un domaine donné correspond une sémantique unique des attributs.

Nous allons prendre deux exemples particuliers pour illustrer ces conflits sémantiques : les relations de composition de type portion / masse (Winston, Chaffin et al. 1987; Markowitz, Nutter et al. 1992)) et la relation de connexion (afférences et efférences³) entre structures. Puis les conflits intra-domaines ou conflits sur des valeurs seront illustrés par d'autres exemples dans un domaine particulier, avant de présenter le mécanisme d'héritage qui permet de résoudre ces conflits par une désignation explicite, à deux niveaux. De telles stratégies de désignation explicite, souvent appelées "point de vue", ont déjà été appliquées dans ROME (Carré, Dekker et al. 1990; Dekker and Carré 1992), FROME (Carré and Geib

³ Les connexions entre les structures du cerveau possèdent généralement un sens privilégié et s'appellent afférences lorsque les signaux arrivent dans la structure considérée et efférences lorsque les signaux sortent de la structure considérée.

1990), OBJLOG (Chouraqui and Dugerdil 1988; Dugerdil 1988; Dugerdil 1991)), SHOOD (Nguyen and Rieu 1991).

En anatomie, le corps strié est composé du noyau caudé et du noyau lenticulaire, lui-même formé du putamen et du pallidus. Au niveau anatomique, le noyau caudé et le noyau lenticulaire sont bien distincts, mais, du point de vue fonctionnel, le noyau caudé et le putamen forment une même unité fonctionnelle appelée striatum ; en outre, ces deux structures sont ontogénétiquement d'origine télencéphalique alors que le pallidus est d'origine diencéphalique. Or, le striatum est principalement identifié dans le domaine neurologique, tandis que le corps-strié et noyau-lenticulaire apparaissent essentiellement dans le domaine anatomique (cf. Figure 14).

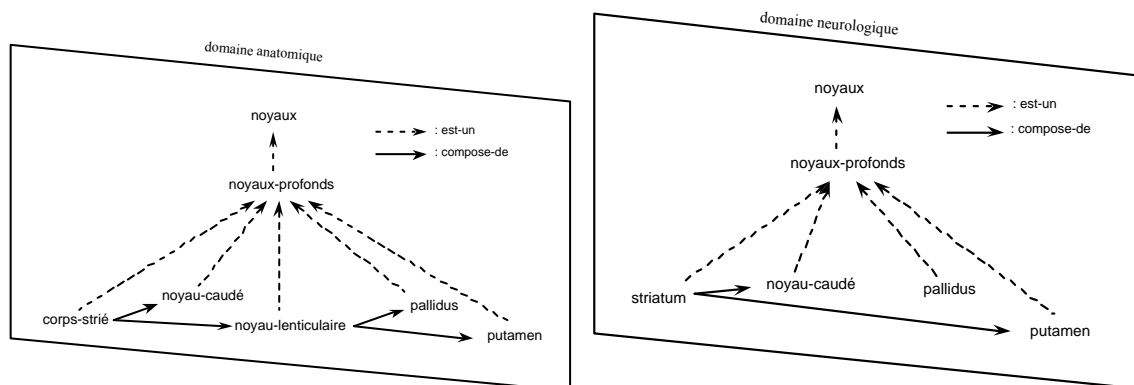


Figure 14: Relation de composition de type portion / masse en fonction des domaines

Dans la figure 3, la relation "composant-de" (inverse de "composé-de") a les mêmes propriétés logiques dans les deux domaines, mais elle n'a pas la même sémantique. En effet, les causes de ces regroupements ou composition ne sont pas les mêmes dans chaque domaine. Par conséquent, pour le noyau-caudé ou le putamen, ainsi que pour tout descendant de ces objets, la valeur de l'attribut "composant-de" n'est pas la même pour chaque domaine ; en neurologie pour le noyau-caudé la valeur est "striatum" et en anatomie la valeur est "corps-strié".

Ainsi, si l'on s'intéresse aux connexions entre le striatum et la substance noire, on dira que putamen et noyau caudé ont, notamment, des afférences provenant de la substance noire (afférences nigro-striées) et ont des efférences qui gagnent, en particulier, cette même substance noire (fibres strio-nigriques). D'un point de vue neurochimique, on pourra préciser que les afférences nigro-striées sont dopaminergiques et ont un rôle inhibiteur sur le striatum ; en outre, ces fibres prennent naissance dans la pars compacta de la substance noire, où se fait la synthèse de la dopamine. De même, on pourra distinguer parmi les fibres strio-nigriques deux types différents : les unes GABA-ergiques ayant un rôle inhibiteur de la substance noire, les autres facilitatrices, utilisant un autre neurotransmetteur, la substance P ; ces fibres se terminent dans une partie de la substance noire, appelée pars reticulata (cf. Figure 15) :

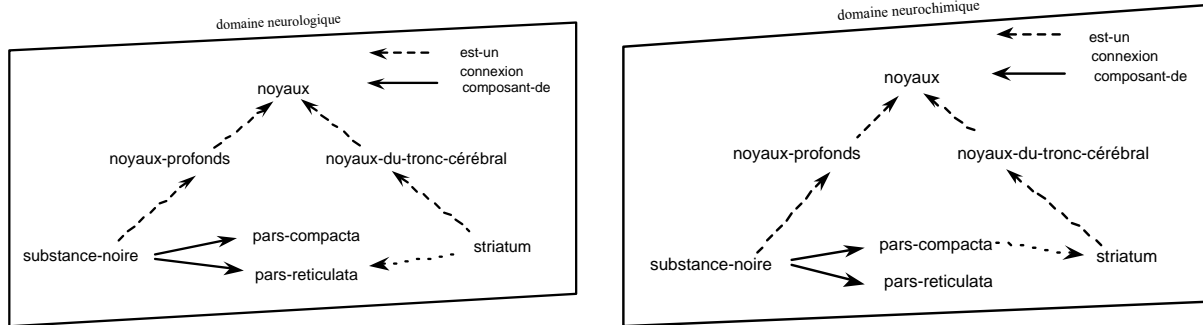


Figure 15: Relations de connexion entre substance-noire et striatum

Dans la figure 15, la relation de connexion n'a pas la même sémantique dans les deux domaines. La valeur de l'attribut "connexion" n'est donc pas la même dans le domaine neurologique (valeur : "pars reticula") et neurochimique (valeur : "pars compacta").

Comme nous avons pu le voir précédemment, la sémantique d'une relation (composition ou connexion) n'est pas la même dans des domaines distincts et possède, évidemment, des valeurs différentes par domaine. En effet, ce ne sont pas les mêmes processus physiques qui sont mis en œuvre dans les différents domaines. En d'autres mots, si l'on devait décrire la sémantique des relations par des traits pour chaque domaine, ces traits seraient très différents. A chaque domaine correspond une sémantique unique pour chaque attribut.

Par contre dans un domaine, l'héritage multiple est autorisé et peut être source de conflit. Ce type de conflit porte sur des valeurs. Il peut être notamment dû à des analyses différentes (des variantes) du domaine et ne modifie pas la sémantique des attributs mais peut modifier la valeur de l'attribut.

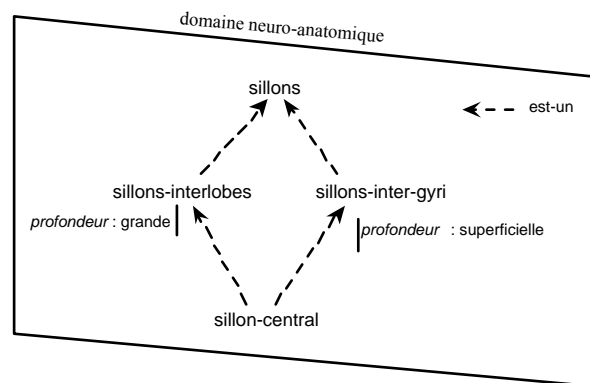


Figure 16: Exemple de conflit inter domaine

Dans l'exemple de la figure 16, le sillon central doit hériter de la valeur "grande" pour l'attribut "profondeur". En effet, le sillon-central sépare des gyri (le gyrus-précentral et le gyrus-post-central) et des lobes (le lobe-frontal et le lobe-pariétal) ; le sillon-central est un sillon-interlobes parce qu'il est profond. L'autre valeur présente dans sillons-inter-gyri n'est pas correcte. Il faut donc pouvoir sélectionner la bonne valeur. Une stratégie globale, comme, par exemple, la linéarisation, privilégierait toujours l'héritage par sillons-interlobes ; mais pour une autre propriété, cette stratégie risque de se révéler illicite.

2.3.2.2.2 L'architecture de la base de connaissances

Dans notre cadre, les perspectives représentent la structure conceptuelle d'un domaine d'expertise particulier. La structure conceptuelle d'un domaine est en fait similaire à une "terminologie". Comme montré par Condamines and Amsili (Condamines and Amsili 1993), la notion de terme peut être vue comme un mot qui recueille des propriétés linguistiques ou comme l'étiquette d'un concept qui regroupe des connaissances sur le "monde". En d'autres termes, le premier réfère à des connaissances lexicales et le second à des connaissances conceptuelles (Rastier 1997). Nous sommes uniquement concernés ici par la seconde interprétation. Elle est également proche de celle des domaines sémantiques utilisés par Pottier (Pottier 1974). Dans un domaine, chaque objet et attribut possède une sémantique unique. Par conséquent, il n'y a pas de conflit sémantique dans un domaine, mais uniquement entre les domaines. Dans un domaine, seuls des conflits de valeurs peuvent se produire. Une telle organisation sépare naturellement les conflits en conflits inter-domaines et intra-domaine qui sont respectivement des conflits sémantiques et de valeurs. Nous allons introduire de manière plus précise la structure de la base de connaissances.

- Soit $D = \{d_1, d_2, \dots, d_n\}$ un ensemble de domaines tel que $d \in D$, $G_d = (X_d, H_d)$, avec X_d l'ensemble des objets du domaine d , où $X_d = X'_d \cup X''_d$, X'_d l'ensemble des classes, X''_d l'ensemble des instances, H_d l'ensemble des liens tels que $H_d = H'_d \cup H''_d$, H'_d l'ensemble des liens "is-a" et H''_d l'ensemble des liens "member-of". Nous pouvons ainsi définir le graphe $G'_d = (X'_d, H'_d)$ qui possède les propriétés suivantes :
- $\forall d_i, d_j \in D, X'_{di} \cap X'_{dj} = \emptyset$ et pour tout $d_i \in D, i = 1, \dots, n, n = |D|$ avec $|D|$ le cardinal de D tel que $|D| > 1, \bigcup_{i=1}^n X'_{di} = X'$. Les ensembles X'_{di} forment une partition de l'ensemble X' (ensemble de classes) et sont, par hypothèse, non vides.
- Un graphe G'_d un graphe orienté sans circuit ; l'héritage multiple est autorisé dans un domaine.
- On peut diviser les relations entre les objets en deux sous-ensembles : le premier contient les relations intra-domaine et le second les relations inter-domaines. Actuellement, il n'existe qu'une relation inter-domaines : la relation d'égalité extensionnelle (coréférence) notée "Eq-Extp" ; elle relie deux classes si et seulement si : soient x et y des classes telles que $x \in X'_{di}$ et $y \in X'_{dj}$, $(Eq-Extp(x,y) \Rightarrow Extp(x) = Extp(y))$. C'est une relation d'équivalence. Il peut être nécessaire de maintenir la cohérence entre les classes reliées par cette relation quand un domaine est modifié ou quand un nouveau domaine est ajouté.
- $\exists d_i, d_j \in D$ tel que $X''_{di} \cap X''_{dj} \neq \emptyset$; pour tout $d_i \in D, i = 1, \dots, n, n = |D|, \bigcup_{i=1}^n X''_{di} = X''$. Une instance peut appartenir à plusieurs domaines mais ce n'est pas une obligation. En effet, certaines entités n'existent que dans un domaine et pas dans

d'autres. Une instance appartenant à une classe x doit aussi appartenir à la classe y si celle-ci est dans autre domaine telle que la relation $\text{Eq-Ext}_p(x, y)$ est satisfaite.

L'organisation de la base de connaissances est donc la suivante (cf. Figure 17) :

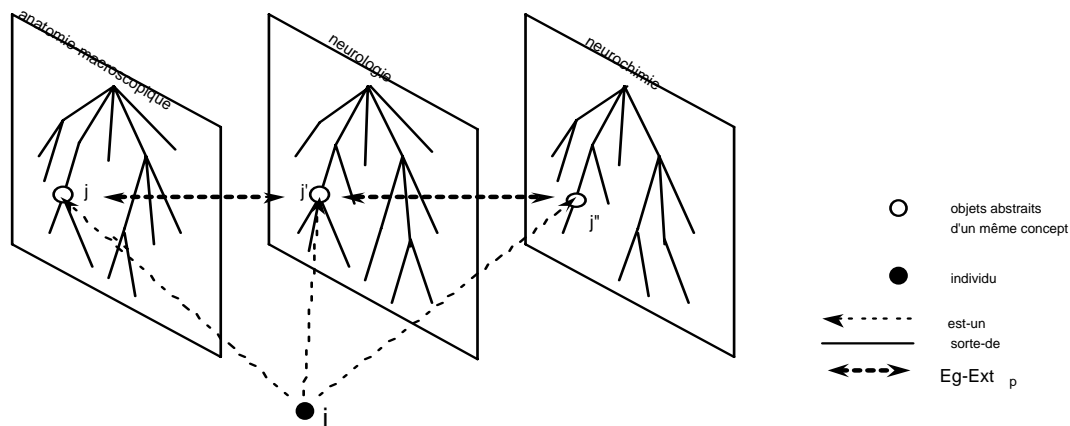


Figure 17 : Organisation de la base de connaissances

Dans la figure 17, l'instance i est reliée aux trois classes j , j' et j'' qui sont extensionnellement égales. Une instance peut être membre de plusieurs classes - incomparables entre elles par la relation \leq_H - par domaine. Chaque domaine possède son propre espace d'attributs qui est séparé de celui des autres. Dans un domaine, si un conflit de valeur se produit, l'objet considéré n'héritera que d'un seul attribut. Le principe d'héritage de toutes les caractéristiques d'un attribut dont son essence est obtenu par défaut ; en fait, tous les attributs nommés A dans un domaine dénote la même entité. En d'autres termes, il n'est pas nécessaire que deux déclarations de l'attribut soient faites dans deux classes reliées par \leq_H pour dénoter la même entité. Par contre, si une instance peut hériter d'un attribut nommé A de plusieurs domaines, elle en héritera d'un par domaine. Il s'agit maintenant de définir le mécanisme d'héritage.

Dans TROPES, une base de connaissances se compose d'un ensemble de concepts disjoints. Chaque concept, présent dans plusieurs domaines d'expertise, est représenté par plusieurs perspectives - des arbres. Les racines de ces arbres sont reliées par un pont - coréférence. Dans la figure 17, les classes j , j' et j'' peuvent jouer le rôle de ces racines. Une perspective d'un concept (sous-graphe de ces descendants) est l'équivalent dans notre cas d'un sous-graphe ayant pour racine une classe d'un domaine - classe reliée à d'autres par une relation d'égalité extensionnelle. Un domaine pourrait contenir plusieurs perspectives disjointes - une par concept présent dans le domaine - plus d'autres classes présentes uniquement dans ce domaine. La notion de perspective ici est en quelque sorte plus générale que celle de TROPES. Comme dans TROPES, le sous-graphe associé à une instance dans une perspective est composé de la hiérarchie de celle-ci et pour les classes de la hiérarchie de la classe dans la perspective considérée.

2.3.2.2.3 Le mécanisme d'héritage

Le mécanisme d'héritage doit satisfaire le masquage, les contraintes associées à la résolution des conflits sémantiques et de valeur tout en prenant en compte l'architecture de la

base de connaissances. Ce mécanisme doit traiter trois cas : les conflits sémantiques sans conflit de valeur, les conflits de valeur sans conflit sémantique et les deux types de conflit simultanément. En cas de conflit sémantique uniquement, la désignation explicite du domaine résout ce conflit en déterminant l'attribut considéré - celui du domaine. En cas de conflit de valeur uniquement, la désignation explicite d'une classe particulière permet de déduire la valeur de l'attribut dans un domaine donné. Lorsque les deux types de conflit existent, il est nécessaire de procéder en deux étapes. La première résout les conflits inter-domaines ou sémantiques et la seconde les conflits de valeur. La première étape est très simple puisqu'elle consiste à désigner un domaine. Pour résoudre la seconde étape, il est nécessaire de réduire l'ensemble de conflit au domaine concerné.

Définition 7 : Soient $d \in D$, A un attribut et α un objet ; l'ensemble de conflit relatif au domaine d , noté $CS(\alpha, A, d)$, est l'ensemble de conflit $CS(\alpha, A)$ réduit au sous-graphe $G_d(\alpha)$.

Autrement dit, les objets qui n'appartiennent pas au graphe G_d sont exclus de $CS(\alpha, A)$. L'ensemble $CS(\alpha, A, d)$ possède les caractéristiques suivantes :

- $\forall \alpha \in X'$ et $\forall d \in D$, $CS(\alpha, A, d) = CS(\alpha, A)$, puisque chaque objet abstrait n'appartient qu'à un unique domaine. Les conflits sémantiques n'ont de sens que pour les instances.
- $\exists \alpha \in X''$ et $\exists d \in D$ de sorte que $CS(\alpha, A, d) \subset CS(\alpha, A)$, puisqu'il existe des objets individuels pouvant appartenir à plusieurs domaines.

Après désignation d'un domaine, la résolution d'un conflit de valeur doit satisfaire le principe 5.

Principe 5 : Pour un objet α et un attribut A et un domaine d , la valeur de l'attribut A ne peut être obtenue qu'à partir d'un objet de $CS(\alpha, A, d)$, $d \in D$.

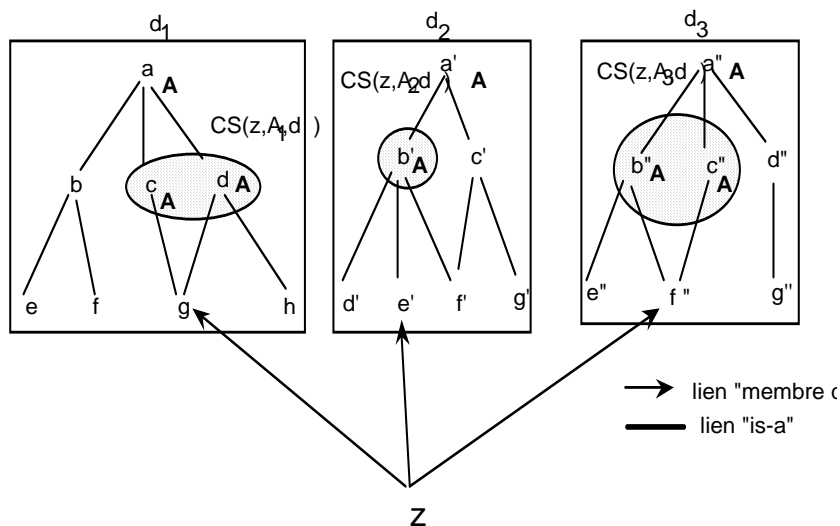


Figure 18 : Exemple de graphe d'héritage avec conflits

Dans l'exemple présenté figure 18, les deux types de conflit sont présents. L'ensemble de conflit $CS(z, A)$ est égal à $\{c, d, b', b'', c''\}$. Le premier niveau de sélection choisit un domaine, par exemple $d1$, qui détermine une sémantique pour l'attribut A . L'ensemble de conflit - conflits de valeur - relativement à $d1$ se réduit à $CS(z, A, d1) = \{c, d\}$. Ce dernier contient deux objets qui possèdent l'attribut A . Le deuxième niveau de sélection nécessite de choisir, implicitement ou explicitement, c ou d pour hériter de A , en les désignant explicitement. Si l'on s'intéresse à $CS(g, A, d1)$, ce dernier est égal à $\{c, d\}$ et il faut donc choisir de nouveau entre les deux (i.e. c et d). Par contre si on désigne le domaine $d2$, il n'y aura pas de conflit de valeur puisque $CS(g, A, d2)$ est égal à $\{b'\}$. Si z possède l'attribut A , il faut préciser s'il s'agit de A par rapport au domaine $d1$, $d2$ ou $d3$. En effet, si z possède A avec la sémantique du domaine $d1$ seule, alors $CS(z, A, d1)$ se réduit à $\{z\}$; par contre $CS(z, A, d3)$ est toujours égal à $\{b'', c''\}$. Pour le second niveau de sélection, le problème majeur est le choix d'une classe permettant de désigner implicitement ou explicitement la valeur attendue. Nous allons aborder ce problème dans la suite.

2.3.2.2.4 Principes de l'algorithme de résolution de conflit

Au vu du mécanisme d'héritage, deux étapes dans l'algorithme de résolution de conflits qui appliquent la stratégie de désignation explicite sont nécessaires. La résolution des conflits sémantiques est simple, puisqu'il suffit de désigner le domaine concerné. Pour les conflits de valeurs, il faut en fait désigner une classe de l'ensemble de conflit dans un domaine. Il suffit donc de désigner ici l'une des classes appartenant à cet ensemble de conflit et y chercher la valeur de l'attribut considéré. Ce procédé est très proche de la notion de point de vue introduite dans ROME. En effet, pour une paire donnée, la résolution du conflit consiste à sélectionner une classe permettant de discriminer les différentes interprétations d'un attribut, la valeur de cet attribut se trouvant dans le sous-graphe associé au point de vue. Toute modification de la hiérarchie considérée ou des déclarations de l'attribut peut entraîner des erreurs si l'on va uniquement chercher la valeur de l'attribut dans la classe désignée. En effet, la déclaration peut avoir été supprimée ou une déclaration plus spécifique faite. Contrairement à la stratégie de linéarisation, il n'est pas possible de prendre en compte tous les types de modifications. Seules celles préservant les propriétés suivantes peuvent être traitées :

1. Pour une paire donnée (x, A) , la classe désignée C appartient toujours à $G_d(x)$,
2. L'ensemble de conflit restreint au sous-graphe associé au point de vue référencé par C pour la paire (x, A) est un singleton, c'est-à-dire que la désignation de C permet toujours de résoudre le conflit
3. De nouveaux conflits de valeurs ne surgissent pas suite aux modifications
4. Il existe bien une valeur pour A dans ce sous-graphe.

Chercher la valeur de A pour x revient donc à faire ces vérifications, puis à parcourir ce sous-graphe en respectant le principe du masquage pour trouver la valeur de A . Le sous-graphe noté $G_d(x, C)$ est en fait la hiérarchie de $G_d(x)$ restreinte aux descendants de C ($Desc(C)$) et à la hiérarchie $G_d(C)$.

L'algorithme consiste en un marquage des classes appartenant à l'ensemble de conflit de la paire considérée et de celles qui appartiennent au sous-graphe associé à la désignation de la

classe C, avec une vérification du cardinal de l'ensemble de conflit restreint au sous-graphe $G_d(x, C)$. Le premier marquage se fait par un parcours en profondeur glouton qui respecte le masquage et l'autre par un marquage des ancêtres de C et des descendants de C. Après satisfaction des contraintes citées précédemment, la valeur de l'attribut se trouve dans la classe ayant les deux marquages (Garlatti, Montabord et al. 1995). Cet algorithme à une complexité linéaire en nombre de noeuds et de liens du graphe considéré.

Nous avons proposé une organisation de la base de connaissance en différents types de connaissances faiblement couplées - connaissances conceptuelles de domaines provenant des disciplines médicales concernées (neurologie, neurochirurgie, neurophysiologie, ...) et des connaissances contextuelles décrivant des tâches munies de leurs domaines ou sous-domaines pertinents, ainsi qu'une organisation des domaines et sous-domaines. La base de connaissances est donc divisée en deux parties : les connaissances de domaine et les connaissances contextuelles. La base de connaissances contextuelles va permettre de réaliser une forme de filtrage pour la recherche d'information fondée sur le mécanisme de navigation de l'hypermédia en déterminant en fonction du contexte les domaines ou sous-domaines pertinents pour l'utilisateur. Il est difficile (voir même impossible) de concevoir une base de connaissance unique qui inclut chaque discipline. Il devient donc nécessaire de séparer les différentes disciplines tout en assurant les « interactions » entre elles. Chaque domaine peut être analysé séparément et indépendamment afin d'obtenir une structure conceptuelle par domaine. Ensuite, il s'agit d'identifier les entités communes à différents domaines – interactions entre les domaines. Nous avons donc proposé une méthodologie de conception pour une base de connaissance intensionnelle composée de plusieurs perspectives quasi-indépendantes appelées domaines et d'un mécanisme de gestion des conflits d'héritage multiple. Les entités communes de différents domaines sont donc reliées par une relation d'égalité extensionnelle. Cette organisation des connaissances en différentes catégories faiblement couplées permet de faciliter la mise à jour des connaissances qui pourra se faire de manière quasi indépendante, mais aussi l'extensibilité de celles-ci puisqu'il devient possible d'ajouter de nouveaux domaines sans remettre en cause les autres.

2.3.3 Conclusion

Nous avons présenté un système d'information sémantique pour le traitement des épilepsies focales. Ce dernier est fondé sur un hypermédia et doit fournir à l'utilisateur les informations pertinentes en fonction du contexte : discipline médicale concernée et tâche à réaliser. Les tâches sont implicites puisqu'elles ne décrivent pas les activités de l'utilisateur mais uniquement les domaines et sous-domaines pertinents pour une tâche donnée. La description des tâches et leur décomposition éventuelle en sous-tâches définit un sous-domaine pertinent parmi les domaines représentés afin de réduire les informations accessibles pour chacune d'elles. Ces domaines ainsi que les connaissances sur les tâches correspondent à une partie des connaissances explicites et partagées de différentes communautés de pratiques – une par discipline. Cet hypermédia est généré dynamiquement à partir des interactions avec l'utilisateur. Les informations sont indexées à partir de ces modèles de domaines. La sélection des informations est réalisée par navigation dans un domaine. Il y a bien réutilisation des informations, mais localement. Il n'y a pas ici de modèle utilisateur ni d'adaptation. La sélection des informations pertinentes est fonction de la discipline et de la tâche à réalisée.

Les travaux réalisés dans le cadre du projet ATLAS ont ensuite permis d'aborder le projet ATLAS-TUTOR qui étudie un système d'enseignement assisté par ordinateur fondé sur le système MR_TUTOR du projet MEDiate et sur l'apport des atlas informatisés et de leurs connaissances pour l'apprentissage du diagnostic en neuroradiologie.

2.4 Le projet ATLAS-TUTOR

Ce projet a pour objet d'étudier l'apport des connaissances anatomiques, ainsi que des connaissances spatiales qualitatives en neuroanatomie pour la conception d'un système tuteur intelligent. Il s'agit d'aider des radiologistes à décrire et interpréter des images de résonance magnétique du cerveau pour leur formation et ainsi offrir une assistance au diagnostic. Le système MR Tutor développé au Sussex a une connaissance médicale limitée, mais fournit une interface multimédia pour l'élicitation des descriptions d'images par un utilisateur. Le but est donc d'étudier une représentation de connaissances permettant au système de répondre à des descriptions incorrectes d'images faites par des "étudiants" à différents niveaux d'expertise. Les experts radiologiste acquièrent une bonne compétence en intégrant des connaissances médicales et pratiques. Cette intégration est un processus actif qui nécessite : a) d'articuler des connaissances médicales ; b) de les appliquer à l'étude des cas – données patients ; c) d'établir des connexions entre ces connaissances médicales et les expériences pratiques. Dans notre cas, il s'agit d'étudier les connexions entre des connaissances anatomiques et spatiales sur les structures du cerveau et des lésions, et la description et l'interprétation des images (Garlatti and Sharples 1998).

2.4.1 Problématique

Une étude sur la formation médicale au Royaume-Uni (Towle 1991) a identifié différents problèmes avec les méthodes courantes d'enseignement et a proposé un programme d'études qui inclut : accentuer l'acquisition de compétences et de savoir-faire plutôt que la communication de faits ; introduire l'apprentissage par résolution de problèmes, l'enseignement à distance et l'évaluation objective et équitable. Les enseignants en radiologie doivent non seulement transformer les méthodes d'enseignement, mais aussi enseigner l'apparence normale et anormale de différentes catégories d'images (différentes modalités), permettre d'accéder à une grande base de cas classés et enseigner une approche cohérente pour le diagnostic des différents cas.

Deux équipes différentes, l'une au Royaume Uni et l'autre en France, abordent ces problèmes en développant, respectivement, un système d'enseignement assisté par ordinateur pour la neuroradiologie des images de résonance magnétique (Sharples, du Boulay et al. 1995; Sharples and Garlatti 1996) et un atlas informatique du cerveau (Montabord, Gibaud et al. 1993). Ces systèmes ont été développés et sont actuellement évalués pour l'usage médical. Un but général de ces deux groupes de recherche est de fournir un environnement qui propose une approche par pédagogie active qui peut s'adapter aux besoins et aux compétences de l'utilisateur et peut couvrir un éventail de tâches en radiologie allant de l'interprétation d'images, au diagnostic et à la planification chirurgicale.

Pour proposer une formation efficace et une aide au diagnostic, le système d'information médical doit avoir une représentation de connaissance profonde du domaine et de l'apprenant qu'il utilise pour produire des séquences d'action d'apprentissage et pour diagnostiquer les

erreurs de celui-ci. Les actions d'apprentissage qu'un enseignant en radiologie effectue peuvent être classées de la manière suivante :

- Instruction : donner la signification d'une limite ou donner une leçon d'introduction, etc.
- Indication : montrer un cas pertinent relativement au diagnostic courant ou montrer des structures anatomiques sur une image, etc.
- Remédiation : détecter et corriger les erreurs, etc.
- Interrogation : poser des questions pour tester les connaissances d'un apprenant, etc.
- Gestion : par exemple présenter un nouveau cas pertinent à résoudre, etc.
- Aide : aider à la navigation ou encourager l'apprenant, etc.

Un système d'information médical fondé sur des connaissances pourrait en théorie offrir toutes ces catégories d'actions d'apprentissage en combinant des connaissances biomédicales et des connaissances en situation (au sujet des cas déjà traités, par exemple), des stratégies d'enseignement fondées sur des règles et un modèle dynamique des connaissances et compétences de l'apprenant. En pratique, les difficultés pour formaliser les connaissances médicales et pour concevoir ce type de système d'information utilisant du multimédia font qu'un faible nombre de systèmes de ce type ont été développés pour la formation en imagerie médicale jusqu'à présent. En dépit de ces difficultés, la combinaison d'un apprentissage à partir de cas et d'une aide au diagnostic fondée sur des connaissances doit permettre de proposer un système d'enseignement flexible et puissant. Une équipe comportant des concepteurs de logiciel, des statisticiens médicaux et des neuroradiologistes développe un système de ce type dans le cadre du projet MEDiate.

2.4.2 Le projet MEDiate

MEDIATE est un projet commun entre l'université du Sussex, l'université De Montfort, à Leicester, l'université de Birmingham et l'institut de neurologie de Londres. Son but est de satisfaire un besoin identifié par les spécialistes en radiologie : une approche plus structurée du diagnostic et de l'apprentissage. Un des résultats du projet est MR-TUTOR : un système d'apprentissage pour aider des radiologistes à interpréter les images MRI (de résonance magnétique) du cerveau, en particulier les images présentant des maladies qui sont connues pour être difficiles à diagnostiquer et différencier.

MR-TUTOR associe un tuteur actif avec un système d'aide adaptatif. Il est fondé sur une approche constructiviste de l'apprentissage dans laquelle l'apprenant acquiert une approche structurée pour décrire les structures anatomiques anormales sur les images en s'engageant dans un processus de diagnostic des cas. L'équipe MEDIATE a conçu un langage structuré de description d'images (IDL) qui est approprié pour une large gamme de séquences d'images⁴.

⁴ Langage qui a été adopté au niveau européen par les neuroradiologistes

Le langage couvre les aspects visuels et les localisations anatomiques des anomalies. Une archive d'environ 1300 cas, entièrement annotés en utilisant IDL munie des informations cliniques et des diagnostics d'experts, forme la base des connaissances situées de MR-TUTOR.

Un apprenant peut consulter les différents cas de la base. Ce dernier ou le système peut sélectionner des séquences d'images pour apprendre en résolvant un problème de diagnostic. L'apprenant identifie et décrit les structures anatomiques anormales dans les images en utilisant la terminologie proposée par IDL et le système fournit des aides pour une interprétation complète et correcte du cas courant. Les termes utilisés dans le langage IDL représente un vocabulaire commun et partagé par tous les neuroradiologistes européens. Il est le résultat d'un consensus au sein du groupe européen des neuroradiologistes et représente donc des connaissances communes explicites de cette communauté de pratiques. L'une des caractéristiques de MR-TUTOR est d'utiliser une analyse des correspondances (Greenacre 1993) pour produire une vue en deux dimensions répartissant les cas dans un nuage de points en fonction de leur apparence et de la localisation des structures anatomiques anormales dans les images pour une maladie donnée. La position des cas dépend d'une mesure de similitude entre eux (en termes de distance entre les points représentant les cas) et de typicalité (distance d'un point au centre de la vue) et permet à l'utilisateur d'accéder ainsi aux images d'archive en agissant sur l'un ou l'autre des points de la vue (cf. figure 19). En regardant les nuages de points, - un par maladie - l'apprenant peut avoir une vue générale (une configuration) de la répartition des cas en fonction de la maladie. Ils peuvent ainsi identifier les cas limites qu'il est difficile de diagnostiquer.

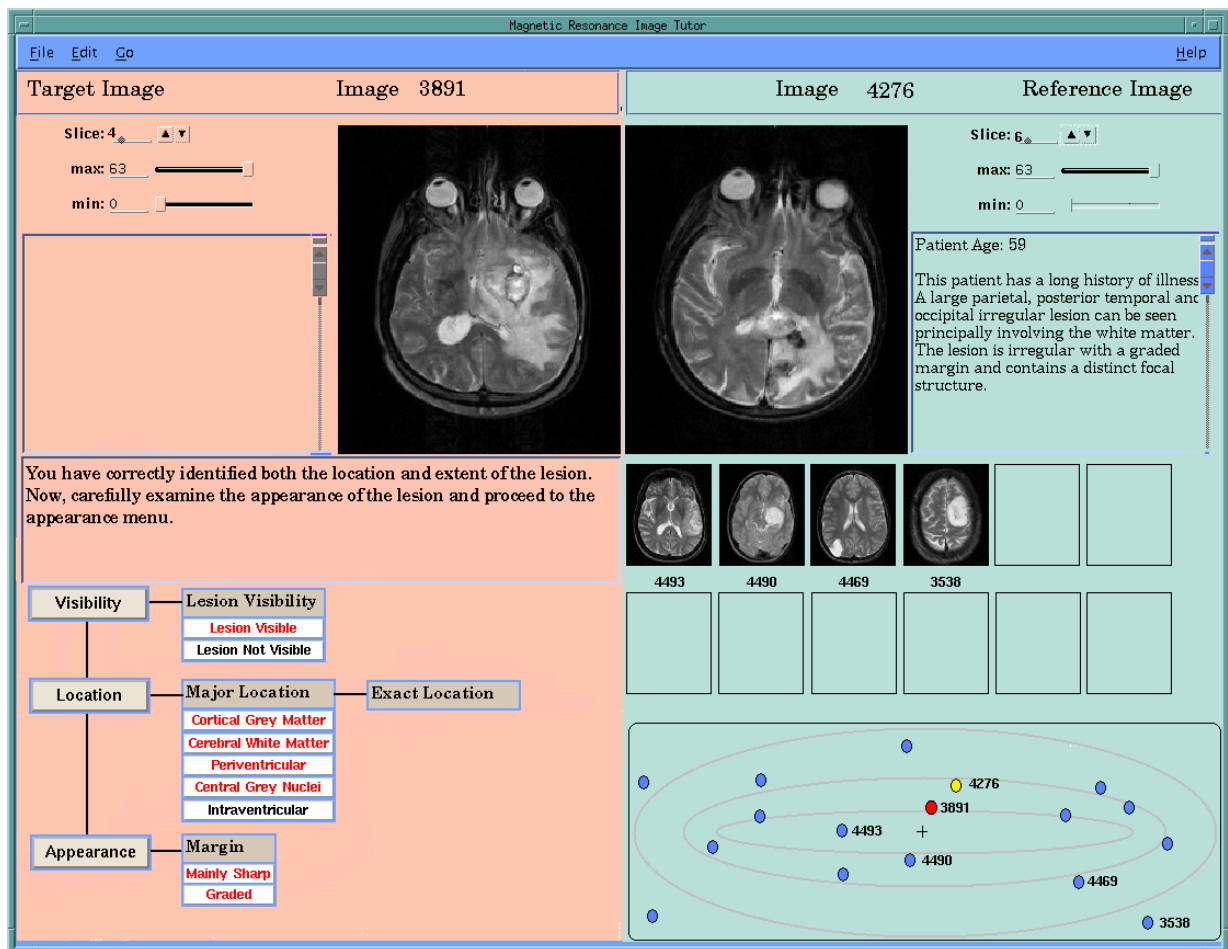


Figure 19 : Interface de MR-TUTOR,

L'exemple de la figure 19 montre deux nuages de points l'un pour le Glioma (points foncés) et l'autre pour l'infarctus (points clairs) à la droite inférieure de la vue.

IDL a été développé pendant environ dix années, impliquant l'acquisition des connaissances d'un neuroradiologiste expert, pour s'assurer qu'il est précis, nécessaire et suffisant pour réaliser le diagnostic de maladies parfois difficilement discernables (du Boulay, Teather et al. 1994). La validité de ces nuages de points a été évaluée en demandant à des neuroradiologistes d'observer un ensemble de cas associés à un nuage de points pour une maladie donnée (gliomas), puis de placer de nouveaux cas dans ce nuage. Les résultats montrent qu'il y a une forte corrélation entre le placement des mêmes cas par les neuroradiologistes expérimentés et par l'algorithme d'analyse des correspondances.

Une évaluation heuristique de MR-TUTOR a été effectuée pour évaluer sa rentabilité et son efficacité générale comme système de formation. L'évaluation a impliqué deux groupes de sujets, l'un composé de personnes expérimentées en conception de système informatique et d'interaction homme/machine et l'autre de neuroradiologistes expérimentés. Sur une échelle de notation de 5 pour la satisfaction globale, l'estimation moyenne du groupe d'interaction homme/machine était 3.96 et celui des radiologistes était de 4.13. Les deux groupes ont souligné que le système était stimulant et facile à utiliser. Les spécialistes d'interaction homme/machine ont signalé que le système est conforme aux bonnes pratiques du domaine et

plusieurs neuroradiologistes ont indiqué qu'il pourrait fournir une aide utile à leur enseignement. Les deux groupes ont suggéré des améliorations mineures au niveau de l'interface et des fonctionnalités.

Une des limitations de MR-TUTOR est son accès insatisfaisant aux connaissances structurées, aux images et aux illustrations de l'anatomie normale et anormale. Lorsque l'apprenant ne précise pas si une ou plusieurs structures sont affectées par une lésion ceci peut se produire parce que :

- L'apprenant a mal interprété les limites de la lésion (peut-être en raison de sa similitude en terme d'apparence avec le tissu normal), ou
- L'apprenant se rend compte de la limite de la lésion mais ne peut pas distinguer ou nommer les structures anatomiques qu'elle affecte.

MR-TUTOR pourrait distinguer ces deux types d'erreurs en demandant à l'apprenant de marquer les limites de la lésion. Si l'apprenant ne peut pas circonscrire exactement la lésion le système pourraient se concentrer sur l'apprentissage de l'identification des lésions. Par contre, si l'étudiant peut délimiter correctement la lésion, le système pourrait en déduire que l'apprenant manque de connaissance en anatomie « normale », et qu'il devrait donc fournir des cours sur l'anatomie afin de palier cette carence. Pour cela, le système doit pouvoir afficher les noms et les zones de l'anatomie affectée par la lésion et pouvoir montrer ces derniers sur les images du cas courant. Pour traiter d'autres erreurs, le système pourrait également avoir des connaissances sur les structures anatomiques adjacentes, leurs tailles, leurs formes et l'apparence radiologique des structures anatomiques affectées et leur aspect par rapport à la lésion. La version courante de MR-TUTOR n'est pas capable de faire un enseignement de ce type.

Plus généralement, pour fournir un tel enseignement, le système doit pouvoir montrer les frontières des structures anatomiques dans différentes séquences d'images pour chaque cas et pour différentes orientations, fournir des informations sur l'histologie, la biochimie et les propriétés fonctionnelles de l'anatomie normale et anormale. Un atlas informatisé devrait pouvoir offrir une telle aide. L'intégration d'un atlas informatisé avec un système de formation en radiologie comme MR-TUTOR devrait donc offrir une plus grande flexibilité et une meilleure formation des neuroradiologistes.

2.4.3 Conception d'un ATLAS-TUTOR

L'étude du système MR-TUTOR du projet MEDiate nous a montré les limites de ce système et les besoins en termes de formation des neuroradiologistes. Nous avons décrit six types d'action d'apprentissage qu'un enseignant en neuroradiologie pourrait utiliser : instruction, indication, remédiation, interrogation, gestion et aide. Un atlas informatique peut apporter les éléments nécessaires à la mise en œuvre de chacun de ces types d'enseignement. Avoir toutes ces actions et ces ressources disponibles permet à un système de type ATLAS-TUTOR de modifier dynamiquement ses actions d'enseignement en fonction des besoins de chaque apprenant. On va maintenant analyser comment la composition des deux systèmes peut proposer toutes les actions d'enseignement nécessaires pour l'aide au diagnostic.

2.4.3.1 L'instruction

Un système de formation peut utiliser un atlas informatique pour enseigner l'anatomie de base. Par exemple, il peut proposer des leçons de rappel, demandant à l'étudiant de délimiter, de nommer ou de montrer les structures anatomiques présentes sur les planches d'atlas. C'est l'action d'enseignement la plus simple à mettre en œuvre, il suffit pour cela d'utiliser les connaissances et les ressources disponibles à partir du système ATLAS et d'y ajouter des interfaces spécifiques.

2.4.3.2 L'indication

Si le système détecte que l'apprenant fait des erreurs en identifiant les régions d'une image ou s'il demande des informations sur des structures normalement connues, alors le système affiche les images du cas courant munies des structures anatomiques pertinentes provenant d'un atlas. Pour cela, il faudrait appliquer un modèle de déformation sur les planches de l'atlas munies de structures anatomiques et y superposer les images de cas. Même si les modèles de déformation ont été prévus pour l'anatomie normale, ils peuvent être efficaces pour des cas ayant des lésions.

2.4.3.3 Remédiation

Une des compétences les plus importantes en neuroradiologie est l'identification des lésions et de leur emplacement par rapport aux structures anatomiques. Il y a quatre types généraux d'erreurs conceptuelles/perceptuelles qu'un apprenant pourrait faire en identifiant et en délimitant une anomalie sur une image :

- Omission : l'apprenant n'identifie pas une structure anatomique couverte par l'anomalie.
- Addition : l'apprenant identifie une structure anatomique qui ne peut pas être détectée comme anormale (mais peut être détectée sur d'autres images ou par d'autres méthodes), ou une structure anatomique qui n'est pas anormale.
- Transformation : l'apprenant décrit l'anomalie dans une position incorrecte (c'est-à-dire une combinaison de l'addition et de l'omission).
- Conformité : l'apprenant fait une erreur de jugement sur les limites d'une anomalie dans une structure anatomique.

Généralement, ces erreurs peuvent avoir les causes suivantes :

- L'apprenant n'est pas capable de détecter l'anomalie ou de délimiter l'anomalie ou d'identifier le type de l'anomalie (connaissances de diagnostic insuffisantes), ou
- L'apprenant connaît les limites de l'anomalie, mais confond l'anatomie couverte par celle-ci (connaissances anatomiques insuffisantes).

Un atlas informatique couplé au système MR-TUTOR permet de distinguer les deux causes d'erreur de l'apprenant. S'il a accès à un atlas superposé aux images du cas courant, le

système peut demander à l'apprenant de dessiner les limites de l'anomalie. Il peut ensuite utiliser l'atlas pour identifier toutes les structures anatomiques envahies ou déplacées par la région délimitée par l'apprenant. L'algorithme peut s'exprimer de la manière suivante :

- Identifiez les limites exactes de toutes les anomalies par un expert et les ajouter aux ressources du système.
- Employez un modèle de déformation pour superposer correctement une image de référence de l'atlas et l'image courante d'un cas.
- Pour chaque anomalie, déterminez quelles sont les structures anatomiques affectées, les limites de celles-ci et le type de changement (invasion ou déplacement).
- Produire une description symbolique des relations spatiales de l'anomalie avec les structures anatomiques qui peuvent ensuite être comparées à la description de l'apprenant et être ensuite utilisée pour guider la remédiation.

Ainsi le système peut distinguer les deux types d'erreur : (1) et (2) ci-dessus et faire une remédiation. Si les limites de l'anomalie dessinées par l'apprenant ne s'apparient pas avec celles fournies par l'expert (en fonction d'un seuil), alors l'apprenant devrait avoir une connaissance de diagnostic insuffisante. Le système peut alors proposer des exercices pour améliorer l'identification et la description des anomalies. Si les limites dessinées par l'apprenant sont approximativement correctes, alors l'apprenant a des connaissances anatomiques insuffisantes et le système peut proposer des conseils autour de la différence entre la description de l'anatomie par l'apprenant et celle détectée en superposant l'atlas dans la zone située autour de l'anomalie. Le système peut également utiliser l'atlas pour mettre en valeur les structures anatomiques mentionnées par l'apprenant, montrer ainsi comment elles sont affectées par l'anomalie et donner quelques explications à l'apprenant.

Il est bien évident qu'il reste de nombreux points à résoudre comme par exemple : le choix des modèles de déformations, l'identification du type (invasion ou déplacement) et des limites de la déformation de l'anatomie provoquée par une anomalie. Comme beaucoup de structures anatomiques n'ont aucune limite claire, il est nécessaire d'établir un seuil de tolérance satisfaisant pour la superposition des images d'atlas et des images de patients et pour la détection d'erreur.

2.4.3.4 Interrogation

Un atlas pourrait intervenir de différentes manières pour cette action :

- Demander à l'apprenant de nommer l'anatomie affectée par une lésion.
- Désigner une région sur une image et demander à l'apprenant de nommer l'anatomie correspondante.
- Demander à l'apprenant de dessiner et de définir les limites de toutes les structures anatomiques visibles sur une image.

- Choisir une région de l'image et demander à l'apprenant de désigner les structures anatomiques (invisibles) qui se trouvent dans cette région.

Si le système est capable d'avoir des modèles de déformation pertinents (c'est-à-dire qui assure une superposition des deux types d'images en respectant les localisations des structures anatomiques), alors il peut aussi proposer les questions suivantes et leurs réponses :

- Demander à l'apprenant la liste des noms des structures anatomiques couvertes par la lésion et la comparer à celle produite par le système.
- Sélectionner une région aléatoire sur une image ou proposer l'une de celles produites à partir de connaissances anatomiques, demander au stagiaire de nommer toutes les structures anatomiques présentes et de les comparer à celle produite par le système.
- Comparer les limites des structures anatomiques dessinées par l'étudiant à celles des planches d'atlas et décrire ou montrer les différences.
- Rechercher des informations sur les gyrus et les noyaux qui sont connus pour se trouver dans une région donnée du cerveau.

2.4.3.5 Gestion

Si le système de formation a accès aux connaissances de l'étudiant en anatomie (dédue de l'enseignement précédent de type interrogation et/ou remédiation) et qu'il a la liste des cas déjà traités par l'apprenant qui sont jugés difficiles par ce dernier, alors le système peut choisir d'autres cas impliquant ces structures anatomiques particulières. Alternativement, si le système a une liste des structures anatomiques qui sont généralement difficiles à identifier ou qui sont régulièrement confondues, alors il pourrait produire une séquence de cas impliquant les structures souvent confondues.

2.4.3.6 Aide

Le système peut offrir un enseignement autonome à l'étudiant en proposant un atlas facilement accessible en ligne, avec une variété d'outils pour consulter l'atlas selon différentes projections, visualiser l'anatomie en trois dimensions, superposer des planches d'atlas à n'importe quelle image de cas proposée par l'apprenant et associer les structures anatomiques à la description de fonctions, par exemple du système visuel ou auditif.

2.4.4 Nouveaux types d'action

Grâce à l'utilisation d'un atlas informatisé, il est possible de proposer de nouveaux types d'actions. Parmi celles-ci, nous allons en décrire deux : diagnostic structuré et remédiation guidée par des hypothèses.

Une approche structurée du diagnostic, fondée sur un protocole systématique et un langage pour décrire les images permet aux neuroradiologistes d'échanger leurs résultats grâce à une terminologie commune et à des référents définis pour chaque terme (du Boulay, Teather et al. 1994). Le langage de description d'image utilisé pour le projet MEDIANE fournit des

ensembles de termes prédéfinis qui sont appropriés au diagnostic. Pour l'enseignement, ce langage a l'avantage pour l'apprenant de proposer un ensemble de termes canoniques muni de définitions, d'exemples et d'un curriculum bien structuré. Un atlas permet d'ajouter, à ces définitions de termes anatomiques, des régions sur des images de référence. Si un apprenant veut connaître l'utilisation de la matière cérébrale blanche, toutes les régions impliquant cette matière cérébrale peuvent être signalées sur un atlas, avec un diagramme de niveau de gris montrant la variation de l'intensité de la matière cérébrale blanche pour une séquence particulière.

La remédiation guidée par des hypothèses est une méthode d'enseignement efficace qui a été incluse dans les systèmes d'enseignement assisté par ordinateur comme le tuteur intelligent SOPHIE pour le dépannage électronique (Brown, Burton et al. 1982). Le fondement de la remédiation guidée par des hypothèses est que si un apprenant donne une réponse incorrecte, alors le tuteur intelligent propose des actions à partir de la différence entre les hypothèses de l'apprenant et la réponse correcte. Par exemple, si un apprenant interprète mal une structure anatomique ou une région de l'anatomie couverte par une lésion, le système pourrait visualiser la région nommée par l'étudiant à partir de l'atlas et la mettre en valeur pour la comparer avec l'image du cas courant.

2.4.4.1 Applications d'un Atlas-Précepteur

Un ATLAS-TUTOR peut proposer de nouveaux types d'actions allant de l'enseignement médical initial à l'aide aux spécialistes. Il peut tenir compte des différents niveaux de connaissance en neuroradiologie à l'aide d'un modèle dynamique de compétences et de connaissances de l'apprenant. Des utilisations possibles d'un tel système sont données ci-dessous :

- Un étudiant en médecine générale pourrait utiliser le système pour un enseignement de base en anatomie du cerveau, avec une introduction générale aux images de résonance magnétique en proposant des images de patients illustrant différentes maladies avec des structures anatomiques superposées.
- On pourrait enseigner à un médecin commençant une formation de spécialiste en neuroradiologie une approche structurée du diagnostic, fondée autour d'une terminologie cohérente pour décrire des images anormales. C'est proche du système MR-TUTOR courant, mais ATLAS-TUTOR peut fournir des conseils sur la signification précise des termes anatomiques en signalant les structures appropriées dans les planches d'atlas. ATLAS-TUTOR permet également de proposer des actions de remédiation en contrôlant le choix des cas pour l'apprentissage.
- Un spécialiste en neuroradiologie pourrait avoir une aide au diagnostic grâce à un environnement intégré qui offre : une bibliothèque des cas décrits avec la terminologie standard ; des représentations imagées et symboliques du cerveau ; un mécanisme de recherche d'information pour les cas, des planches d'atlas et des diagnostics ; accès à la littérature en ligne ; système d'aide à la décision tel que des modèles de déformation pour aider à l'interprétation des images. Un tel environnement permet au spécialiste d'établir les relations entre les connaissances en situation (les cas) et les connaissances médicales d'une manière efficace pour faire son diagnostic.

Pour toutes ces applications, certains problèmes pratiques doivent être surmontés. Nous indiquons ci-dessous certaines de ces questions qui doivent être abordées dans la conception d'un ATLAS-TUTOR performant.

2.4.5 Problèmes de conception

La combinaison d'un atlas fondé sur des connaissances du cerveau avec un système de formation pour la neuroradiologie est une tâche complexe, comportant la formalisation de la connaissance neuroradiologique, les systèmes d'enseignement assisté par ordinateur et le génie logiciel. Dans cette partie, nous discutons de quelques enjeux concernant l'intégration des deux systèmes : un cadre unifié pour la connaissance, des techniques de raisonnement spatial et l'usage des modèles de déformation pour annoter des images de patients.

2.4.5.1 Un cadre unifié pour la connaissance

MR-TUTOR et ATLAS contiennent tous les deux des représentations structurées de la connaissance anatomique. L'équipe d'atlas s'est concentrée sur la terminologie pour décrire l'anatomie normale et l'équipe MÉDIATE sur un langage pour décrire l'aspect anormal. Celles-ci devraient être mises en commun, pour couvrir l'anatomie normale et anormale, à chaque entité est associé un terme standardisé et elle est placée dans un système de classification qui décrit la hiérarchie des classes, des hiérarchies de composition, des relations spatiales (" se touche », « se chevauche », « ci-dessous », « ci-dessus et gauche », etc.), des attributs (" forme », « taille » etc.) et des liens avec les ressources (planches d'atlas, images de patients, données cliniques, etc.). La connaissance doit être anatomiquement précise et appropriée pour l'interprétation en neuroradiologie. Elle devrait différencier les anomalies qui peuvent être vues sur une image et celles qui sont associées aux cas (ils ont été détectés sur une autre image ou par d'autres méthodes) mais ne sont pas visibles. Elle doit également être accessible par l'ordinateur, afin d'être utilisée par le système pour la remédiation.

2.4.5.2 Enregistrement des planches d'atlas sur des images de patients

Pour la formation à base de cas, le système a besoin d'une méthode pour annoter les images de patients avec les limites et les noms de l'anatomie normale et anormale. Faire ceci à la main pour chaque image prend du temps et nous examinons ici des méthodes pour automatiser ou au moins aider les utilisateurs dans ce processus.

Il y a un certain nombre de techniques semi-automatiques pour déformer des planches d'atlas pour le recalage des images de patients. Collins (Collins, Holmes et autres. 1995) décrit un algorithme pour la déformation et le recalage entièrement automatisés. Son approche est de fournir un ensemble de planche d'atlas de référence sous forme d'ensemble de coupes de cerveau normal qui ont été étiquetées avec les noms et les limites de ses structures anatomiques. L'algorithme compare, en trois dimensions, des volumes ayant des positions équivalentes dans les images de patients et les images de référence. Là où l'algorithme détecte une différence dans l'intensité des voxels, indiquant une différence en termes d'anatomie, il déforme l'image d'atlas en fonction de celle du patient. Par une série d'étapes itératives, les volumes globaux en premier et ensuite les volumes locaux de l'image de référence sont recalés. Actuellement, l'algorithme de Collins est limité aux images normales et ne peut pas produire de recalage précis des régions périphériques telles que les gyri et les sillons où il y a

une grande variabilité entre les différents patients. L'algorithme pourrait être amélioré en déterminant une plage d'images de référence pour couvrir des variations qualitatives de l'anatomie.

Bien que Collins ait démontré que cet algorithme peut être appliqué à toute une gamme d'images normales, l'étendre aux séquences d'images contenant des anomalies n'est pas facile. En plus du problème général de sélectionner et de recalcr une planche d'atlas avec une image de patient, une lésion peut envahir ou déplacer l'anatomie environnante, produisant des changements de l'intensité et des déformations de la structure anatomique. L'algorithme de Collins pourrait être étendu avec un procédé additionnel qui déforme la région entourant la lésion, en fonction de la taille de la lésion et du fait que la lésion est connue pour envahir ou comprimer l'anatomie normale. Ceci exigerait l'intervention d'un neuroradiologiste expert pour évaluer le type de lésion et dans certains cas accomplir la déformation nécessaire pour assurer un recalage précis.

2.4.5.3 Raisonnement spatial

Le raisonnement spatial peut jouer deux rôles différents dans un système de formation de type ATLAS-TUTOR : explication et indexation. Si un apprenant ne détecte pas l'anomalie ou comprend mal ses limites (c'est-à-dire manque de connaissance de diagnostic), ou connaît les limites de l'anomalie mais fait des erreurs d'anatomie (c'est-à-dire manque de connaissance anatomique), une action d'enseignement de remédiation pourrait montrer les limites de l'anomalie et recalcr les planches pertinentes de l'atlas avec l'image du patient. Le système a pu également donner des descriptions écrites des relations spatiales entre les structures anatomiques ou entre l'anatomie et les anomalies. Réciproquement, les descriptions écrites pourraient être utilisées comme termes d'index pour construire une requête afin de rechercher des anomalies (par exemple « trouvez toutes les lésions dans l'hémisphère gauche et dans la matière blanche mais pas dans la matière grise non corticale »).

Pour fournir une explication ou répondre à une question, la base de connaissances aurait besoin du raisonnement spatial fondé sur les relations primitives de la topologie (" inclut », « recouvre », « disjoint de », etc.) et de l'orientation (" au-dessus de », « gauche-de ») entre les structures connues. Celles-ci peuvent être combinées en des relations de plus haut niveau par des opérateurs de composition, de disjonction et de conjonction (McNamara 1986; Cohn, Randell et al. 1993; Hernandez 1994).

La tâche d'annotation de chaque image avec des relations spatiales pourrait être automatisée sachant que les structures anatomiques normales et anormales peuvent être décrites et délimitées par un algorithme de déformation. Elle devrait non seulement identifier les limites des structures significatives pour le diagnostic, mais produire aussi des descriptions symboliques de leur forme et leurs positions relatives.

2.4.6 Conclusion

Une nouvelle approche de l'enseignement assisté par ordinateur en neuroradiologie a été proposée. Cette dernière associe un système tuteur intelligent MR-TUTOR avec un atlas informatisé muni de connaissances anatomiques. Un ATLAS-TUTOR permet de proposer un enseignement assisté par ordinateur et une aide au diagnostic dans un même environnement. Il

fournit un enseignement à la demande au travers d'un environnement de multimédia orienté utilisateur et propose une approche structurée du diagnostic en neuroradiologie, complétée par des outils aidant à définir et à identifier des structures anatomiques. Ce système peut être utilisé pour la formation initiale, mais aussi pour des experts en neuroradiologie.

Concevoir un système de formation ATLAS-TUTOR exige davantage de recherche sur des modèles de déformation, de représentation de connaissance et d'interaction humaine/machine. Il est nécessaire de concevoir des stratégies d'enseignement, des modèles dynamiques d'étudiant, et des outils pour le raisonnement spatial. La connaissance conceptuelle sur l'aspect des images anormales, en termes de types de lésion, des formes géométriques, des intensités et des signes associés doit être acquise. Ces connaissances doivent être complétées par la connaissance biomédicale. Les modèles de déformation actuels sont uniquement appropriés aux images de cerveaux normaux, mais nous croyons qu'ils pourraient être adaptés pour des images montrant des anomalies - une étape vers l'interprétation partiellement automatisée d'images. Ces sources de connaissance doivent être composées pour déterminer des actions d'enseignement pertinentes en fonction des connaissances de l'étudiant et du contexte.

2.5 Conclusion générale

Pour cette première étape de l'activité de recherche, les tâches/objectifs des utilisateurs sont pris en compte de manière implicite - il n'y a pas de modèle utilisateur. Les deux applications sont dédiées à des activités d'expertises pour lesquels le processus de « résolution de problème » est entièrement contrôlé par l'utilisateur – le médecin ou le neuroradiologiste. Le système d'information propose donc un mécanisme de navigation contextuelle pour la recherche et l'accès aux informations / connaissances pertinentes et aux outils de visualisation en garantissant la cohérence entre toutes les sources d'information et de connaissances. Il permet aussi l'articulation des connaissances biomédicales et des connaissances en situation. Le contexte prend en compte la discipline médicale ou la tâche du médecin qui détermine une sous-discipline particulière. Cette notion de contexte est malgré tout assez limitée car elle ne permet pas de filtrer efficacement les informations et connaissances pertinentes. Le médecin est donc obligé de sélectionner ces dernières par navigation – en construisant implicitement une requête.

Dans l'étape suivante, la prise en compte de l'adaptation permet de mieux appréhender les besoins utilisateur – donc le facteur humain. Il devient notamment possible de mieux filtrer les informations. Les besoins utilisateur sont représentés explicitement à l'aide d'un modèle utilisateur et le système d'information adaptatif est fondé sur le concept d'hypermédia adaptatif / personnalisable. Le système d'information est capable de prendre en charge un plus grand nombre d'activités qui ne sont plus des tâches d'expertises.

2.6 Crédits

Les études menées dans le cadre du projet ATLAS doivent beaucoup à l'équipe SIM de l'hôpital Pontchaillou et plus particulièrement à M. Bernard Gibaud et M. Christian Barillot, au professeur Scarabin et toute son équipe médicale travaillant notamment sur les épilepsies focales, à Melle Montabord qui a effectuée sa thèse dans le cadre de ce projet et qui a eu la lourde tâche d'acquérir les connaissances médicales, d'associer le traitement d'images et

l'intelligence artificielle et les hypermédia et enfin à Ioannis Kanellos avec qui j'ai co-piloté ce projet et aussi beaucoup partagé et à tous les stagiaires qui ont participé. Pour ATLAS-TUTOR, cette étude n'aurait jamais pu se faire sans l'apport de Ben Du Boulay et de toute son équipe de l'Université du Sussex et tout particulièrement de Mike Sharples, du professeur Du Boulay de l'institut de neuroradiologie de Londres et de l'équipe de l'université de Monfort à Leicester. Plus globalement, ces activités ont profité de toutes les discussions et échanges qui ont eut lieu dans le groupe qui « organisait » la conférence LMO et plus particulièrement Michel Habib, M. Roland Ducournau, Amédéo Napoli, Jérôme Euzenat.

2.7 Bibliographie

- Attardi, G. and M. Simi (1985). Metalanguage and Reasoning Across Viewpoints. Advances in Artificial Intelligence. T. O'Shea, Elseviers Science Publishers B.V.: 413-422.
- Bajcsy, R. and S. Kovacic (1989). "Multiresolution Elastic Matching." CVGIP 46: 1-21.
- Barsalou, L. W. and D. L. Medin (1986). "Concepts: fixed definitions or dynamic context-dependent representations?" Cahiers de psychologie cognitive(6): 187-202.
- Bloom, F. E. (1990). Databases of brain information. Three-Dimensional Neuroimaging. A. W. Toga. New York, Raven Press: 273-295.
- Bobrow, D. G., R. Kaplan, et al. (1977). "GUS, a frame-driven dialog system." Artificial Intelligence 8: 155-173.
- Bobrow, D. G. and M. S. Stefik (1981). The LOOPS Manual, Xerox PARC.
- Bobrow, D. G. and T. Winograd (1977). "An overview of KRL, a Knowledge Representation Language." Cognitive Science 1: 3-46.
- Bohm, C., T. Greitz, et al. (1991). "Specification and selection of regions of interest (ROIs) in a computerized atlas." Journal of cerebral blood flow and metabolism(11): A64-A68.
- Bookstein, F. L. (1991). Thin-Plate Splines And The Atlas Problem For Biomedical Images. Information Processing in Medical Imaging. A. C. F. Colchester and D. J. Hawkes. Berlin, Springer-Verlag: 326-342.
- Borning, A. H. (1986). Classes versus Prototypes in Object-Oriented Languages. IEEE/ACM Fall Joint Conference.
- Boshuizen, H. P. A. and H. G. Schmidt (1993). "On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices." Cognitive science 16: 153-184.
- Brachman, R. J. (1985). I lied about trees or defaults in knowledge representation. AI Magazine: 80-93.
- Brown, J. S., R. R. Burton, et al. (1982). Pedagogical, Natural Language, and Knowledge Engineering Techniques in SOPHIE I, II and III. Intelligent Tutoring Systems. D. H. Sleeman and J. S. Brown. London, Academic Press: 227-282.
- Carré, B. and G. Comyn (1988). On Multiple Classification, Points of View and Object Evolution. Artificial Intelligence and Cognitive Sciences. J. Demongeot, T. Herve, V. Rialle and C. Roche, Manchester University Press: 49-62.
- Carré, B. and G. Comyn (1989). ROME, Multiple and evolutive object representation. Proceedings of the Fifth AIICSR conference, Czechoslovakia, North Holland.
- Carré, B., L. Dekker, et al. (1990). Multiple and Evolutive Representation in the ROME Language, Université de Lille.
- Carré, B. and J.-M. Geib (1990). The Point of View Notion for Multiple Inheritance. ECOOP/OOPSLA, ACM Sigplan Notices.

- Chaffin, R. (1981). "Context effects in the categorization task." The Journal of General Psychology(104): 293-302.
- Chouraqui, E. and P. Dugerdil (1988). Conflict solving in a frame-like multiple inheritance system. European conference on Artificial Intelligence.
- Clayton, B. (1985). ART, Automated Reasoning Tool, Inference Corporation, Los Angeles, California.
- Cohn, A. G. (1993). Modal and Non Modal Qualitative Spatial Logics. Workshop on Spatial and Temporal Reasoning at IJCAI93, Chambéry.
- Cohn, A. G. and N. M. Gotts (1994). A Theory of Spatial Regions with Undetermined Boundaries. Topological Foundations of Cognitive Science. C. Eschenbach, C. Habel and B. Smith.
- Cohn, A. G., D. A. Randell, et al. (1994). Taxonomies of Logically Defined Qualitative Spatial Relations. Formal Ontology in Conceptual Analysis and Knowledge Representation. N. Guarino and R. Poli, Kluwer.
- Cohn, A. G., D. A. Randell, et al. (1993). Qualitative Spatial Reasoning and Representation. QUARDET: Qualitative Reasoning and Decision Technologies, Barcelona.
- Condamines, A. and P. Amsili (1993). Terminology between language and Knowledge: An Example of Terminological knowledge Base. Terminology and Knowledge Engineering, Cologne.
- Cui, Z., A. G. Cohn, et al. (1993). Qualitative and Topological Relationships in Spatial Databases. Lecture Notes in Computer Science No. 692. D. Abel and B. C. Ooi. Berlin, Springer Verlag: 296-315.
- Davis, H. E. (1987). VIEWS: multiple perspectives and structured objects in a knowledge representation language, MIT.
- Dekker, L. and B. Carré (1992). Multiple and Dynamic Representation of Frames with Points of View in FROME. RePrésentation Objets, La grande Motte.
- DeMichiel, L. G. and R. P. Gabriel (1987). The common Lisp Object System: an Overview. ECOOP.
- Downs, R. M. and D. Stea, Eds. (1973). Image and Environnement: Cognitive mapping and Spatial behaviour. Chicago, Aldine.
- du Boulay, G. H., B. A. Teather, et al. (1994). Standard Terminology for MR Image Description. XV Symposium Neuroradiologicum, Kumamoto.
- Ducournau, R. and M. Habib (1989). "Un mécanisme d'héritage dans les langages à objets." Techniques et Sciences Informatiques 8(1): 42-62.
- Ducournau, R., M. Habib, et al. (1995). "L'héritage Multiple dans tous ses états." Technique Scientifique Informatique 14(3).
- Dugerdil, P. (1986). A propos des mécanismes d'héritage dans un langage orienté objet. CIIAM 86 Intelligence Artificielle, Marseille.
- Dugerdil, P. (1988). Contribution à l'étude de la représentation des connaissances fondée sur les objets, Université d'Aix Marseille II.
- Dugerdil, P. (1991). Inheritance mechanisms in the Objlog Language: multiple selective and multiple vertical with points of view. Inheritance Hierarchies in Knowledge Representation and Programming Languages. M. Lenzerini, D. Nardi and M. Simi. Viarregio, John Wiley & Sons Ltd: 245-256.
- Evans, A. C., D. L. Collins, et al. (1992). "An MRI-based stereotactic atlas from 250 young normal subjects." Soc. Neurosci. Abstr(18): 408.
- Evans, A. C., T. S. Marrett, et al. (1991). "Mri-Pet Correlative Analysis Using A Volume Of Interest (Voi) Atlas." J. of Cerebral Blood Flow and Metabolism 11(2): A69-A78.

- Ferber, J. (1988). Coreferentiality: The key to an Intentional Theory of Object-Oriented Knowledge Representation. Artificial Intelligence and Cognitive Science. J. Demongeot, T. Herve, V. Rialle and C. Roche, Manchester University Press. Chap 6.
- Ferber, J. and P. Volle (1988). Using coreference in object oriented representations. European Conference of Artificial Intelligence.
- Fox, P. T., J. S. Perlmutter, et al. (1985). "A Stereotactic Method of Anatomical Localization for Positron Emission Tomography." J. of Computer Assisted Tomography 9(1): 141-153.
- Freksa, C. (1991). Qualitative spatial reasoning. Cognitive and Linguistic Aspects of Geographic Space. D. M. Mark and A. U. Frank. Dordrecht, Kluwer: 361-372.
- Friston, K. J., R. E. Passingham, et al. (1989). "Localization in PET images: Direct Fitting of the Intercommissural (AC-PC) Line." J. Cereb. Blood Flow Metab 9(1): 141-153.
- Garlatti, S., I. Kanellos, et al. (1994). YAKA : un système d'information intensionnel sur le cerveau humain. Informatique des Organisations et des systèmes d'information et de décision, Aix en provence, Inforsid.
- Garlatti, S., E. Montabord, et al. (1995). Knowledge Base Design in Solving Complex Problems. Brest, Télécom Bretagne.
- Garlatti, S. and M. Sharples (1998). "The Use of a Computerized Brain Atlas to Support Knowledge-based Training in Radiology." Artificial intelligence in Medecine 13: 181-205.
- Gee, J. C., C. Barillot, et al. (1994). Matching Structural Images of the Human Brain using Statistical and Geometrical Image Features. SPIE, Visualization in Biomedical Computing, SPIE.
- Goldstein, I. P. and D. G. Bobrow (1980). Extending Object Oriented Programming in Smalltalk. The lisp Conference, Stanford University.
- Greenacre, M. J. (1993). Correspondence Analysis in Practice, Academic Press.
- Greitz, T., C. Bohm, et al. (1991). "A computerized brain atlas: construction, anatomical content, and some applications." Journal of Computer Assisted Tomography 15(1): 26-38.
- Guha, R. V. (1990). Contexts in Cyc, Stanford University, Austin Texas.
- Guha, R. V. (1991). Contexts: A Formalization and Some Applications, Standford.
- Guha, R. V. and D. B. Lenat (1990). Cyc: A Midterm Report. AI Magazine, 3: 33-59.
- Hart, R. A. and G. T. Moore (1973). The development of spatial cognition: a review. Image and Environnement: Cognitive Mapping and Spatial Behaviour. R. M. Downs and D. Stea. Chicago, Aldine: 111-143.
- Hernandez, D. (1993). Reasoning with qualitative representations: Exploiting the Structure of Space. QUARDET, III IMACS international Workshop on Qualitative Reasoning and decision technologies, Barcelona.
- Hernandez, D. (1994). Qualitative Representation of spatial Knowledge. Berlin, Springer verlag.
- Hernandez, D. and K. Zimmermann (1993). Default reasoning and the qualitative representation of spatial knowledge. defaults and prototypes - Non-monotonic Reasoning for Language knowledge Processing. H. e. al, Springer verlag.
- Höhne, K. H., M. Bomans, et al. (1992). "A volume based anatomical atlas." IEEE Comp. Graphics & Appl. 12(9): 72-78.
- Keyserlingk, D. G. v., K. Niemann, et al. (1988). "A quantitative approach to spatial variation of human cerebral sulci." Acta Anatomica(131): 127-131.

- Keyserlingk, D. G. v., K. Niemann, et al. (1985). "A new method in computer assisted imaging in neuroanatomy." Acta Anatomica(123): 240-246.
- Knudsen, J. L. (1988). Name collision in multiple classification hierarchies. ECOOP.
- Lehmann, E. D., D. J. Hawkes, et al. (1991). "Computer aided interpretation of SPECT images of the brain using an MRI derived 3D neuro-anatomical atlas." Medical Informatics 16(2): 151-166.
- Marino, O. (1989). Classification d'objets dans un modèle multi-points de vue, Laboratoire ARTEMIS/IMAG, INPG.
- Marino, O. (1991). Classification d'objets composites dans un système de représentation des connaissances multi-point de vue. RFIA.
- Marino, O. (1993). Raisonnement classificatoire dans une représentation objets multi-points de vue, Université Joseph Fourier, Grenoble.
- Marino, O., F. Rechenmann, et al. (1990). Multiple perspectives and classification mechanism in object-oriented representation. European Conference on Artificial Intelligence.
- Mark, D. M. and A. U. Frank, Eds. (1991). Cognitive and Linguistic Aspects of Geographic Space. NATO Advanced Studies Institute. Dordrecht, Kluwer.
- Markowitz, J. A., J. T. Nutter, et al. (1992). "Beyond is-a and Part-Whole : More Semantic Network Links." Computers Math. Applic 23(6-9): 377-390.
- Martin, W. A. (1979). Descriptions and specialization of concepts. Artificial Intelligence, an MIT perspective. P. Winston and R. H. Brown, The MIT Press. 1: 377-419.
- Mazziotta, J. C., A. W. Toga, et al. (1993). A Probabilistic Reference System for the Human Brain, Application to the Human Brain Project: Phase I.
- McNamara, T. P. (1986). "Mental Representation of Spatial Relations." Cognitive Psychology(18): 87-121.
- Medin, D. L. and E. J. Shoben (1988). "Context and structure in conceptual combination." Cognitive psychology(20): 158-190.
- Meyer, B. (1988). Object-oriented software construction, Prentice Hall.
- Minski, M. (1975). A framework for representing knowledge. The psychology of computer vision. P. Winston. New York, McGraw Hill.
- Mittal, V. O. and C. L. Paris (1993). Context: Identifying its elements from the communication point of view. Workshop on "Using knowledge in its Context", IJCAI-93, Chambéry.
- Montabord, E. (1996). Base de connaissances, hypermédia et contexte, Rennes 1.
- Montabord, E., S. Garlatti, et al. (1994). Interaction between a hypermedia and a knowledge base management system about the brain. neural networks and Expert Systems in Medicine and Healthcare, Plymouth, University of Plymouth.
- Montabord, E., B. Gibaud, et al. (1995). HYPER-YAKA : base de connaissances et hypermédia sensible au contexte. Langage et Modèles à Objets, Nancy.
- Montabord, E., B. Gibaud, et al. (1993). An hypermedia system to manage anatomical knowledge about brain. Computer Assisted Radiology, Berlin, Springer-Verlag.
- Natarajan, K., M. G. Cawley, et al. (1991). "A knowledge based system paradigm for automatic interpretation of CT scan." Medical informatics 16(2): 167-181.
- Nguyen, G. T. and D. Rieu (1991). Database Issues in Object Oriented Design. Tools'91, Paris.
- Niggemann, J. (1990). "Analysis and Representation of Anatomical Knowledge." Applied Art. Intell.(4): 309-336.
- Ono, M., S. Kubik, et al. (1991). Atlas of the cerebral sulci. Stuttgart, Georg Thieme Verlag.
- Pottier, B. (1974). Linguistique générale, théorie et description, Klincksieck.

- Randell, D. A., Z. Cui, et al. (1992). A Spatial Logic based on Regions and Connection. Int. Conf on Knowledge Representation and Reasoning, Boston.
- Rastier, F. (1997). "Représentation du contenu lexical." Langages, sémantique et intelligence artificielle(87).
- Robinson, G. P., A. C. F. Colchester, et al. (1993). Model-Based Recognition of Anatomical Objects from Medical Images. Lecture Notes in Comp. Science: Information Proc. in Med. Imag. H. Barrett and A. Gmitro, Springer-Verlag. 687: 197-211.
- Roth, E. M. and E. J. Shoben (1983). "The effect of context on structure of categories." Cognitive Psychology(15): 346-378.
- Schaltenbrand, G. and W. Wahren (1977). Atlas for stereotaxy of the human brain. Stuttgart, Georg Thieme Verlag.
- Seitz, R. J., C. Bohm, et al. (1990). "Accuracy and precision of the computerized brain atlas programme for localization and quantification in Positron Emission Tomography." Journal of cerebral blood flow and metabolism(10): 443-457.
- Sharples, M., B. du Boulay, et al. (1995). The MR Tutor: Computer-based Training and Professional Practice. World Conference on Artificial Intelligence and Education (AI-ED '95), Washington DC.
- Sharples, M. and S. Garlatti (1996). The Use of Computerised Brain Atlas to Support Case-Based Radiology Training Systems, Télécom Bretagne.
- Simmons, R. and R. Davis (1993). The Role of Knowledge and Representation in Problem Solving. Second Generation Expert Systems. J. M. David, J. P. Krivine and R. Simmons, Springer Verlag: 27-45.
- Sokolowska, E. and J. A. Newell (1986). "Multi-Layered Image Representation: Structure and Application in Recognition of Parts of Brain Anatomy." Pattern Recognition Letters(4): 223-230.
- Staemmler, M., E. Claridge, et al. (1993). "SAMMIE - Software applied to multimodal images and education". IMAC 93, Berlin.
- Stroustrup, B. (1986). Multiple inheritance for C++. EUUG.
- Szikla, G., G. Bouvier, et al. (1977). Angiography of the human cortex. Berlin, Heidelberg, New York, Springer Verlag.
- Talairach, J. and P. Tournoux (1988). Co-Planar Stereotactic Atlas Of The Human Brain. Stuttgart, Georg Thieme Verlag.
- Toga, A. W. (1991). "A Three-Dimensional atlas of structure/function relationships." Journal of chemical neuroanatomy(4): 313-318.
- Towle, A. (1991). Critical Thinking: The Future of Undergraduate Medical Education. London, King's Fund Centre.
- Waern, Y. (1986). Cognitive Aspects of Computer Supported tasks. New York, Wiley.
- Winston, M. E., R. Chaffin, et al. (1987). "A Taxonomy of Part-Whole Relations." Cognitive Science(11): 417-444.
- Wolinski, F. and F. Perrot (1991). Representation of Complex Objects: Multiple Facts with Part-Whole Hierarchies. ECOOP.
- Woods, W. A. (1975). What's in a link: Foundations for Semantics Networks. Representation and Understanding: Studies in Cognitive Science. D. G. Bobrow and A. Collins. New York, Academic Press.
- Zachmann, H. H. (1991). Interpretation of cranial MR images using a digital atlas of the human head. CAR 91, Berlin, Springer Verlag.

Chapitre 3. Un système d'information adaptatif pour la navigation en mer

3.1 Préambule

Cette seconde étape se distingue principalement de la précédente par la prise en compte de l'adaptation aux besoins utilisateur. L'adaptation aux besoins utilisateur est un raffinement de la notion de situation courante de l'étape précédente. En effet, la situation courante est notamment « déterminée » par un modèle de l'utilisateur, un modèle d'adaptation et des stratégies d'adaptation. Les tâches utilisateur n'étant plus des tâches d'expertise, Il devient possible d'accroître l'automatisation dans le système d'information. La problématique principale devient donc celle de la conception d'un système d'information sémantique et adaptatif. Le système d'information doit prendre en charge une part plus importante : des mécanismes de recherche d'information, une structure organisant l'accès aux informations pertinentes en fonction des besoins utilisateur et des mécanismes d'adaptation – filtrage des informations pertinentes. La méthodologie de conception d'un tel système d'information doit donc être revue. En effet, le système d'information est maintenant accessible sur internet au travers d'un navigateur et il doit utiliser de nouvelles catégories de connaissances – un modèle utilisateur, un modèle d'adaptation et un mécanisme d'adaptation.

Dans cette étape, le projet SWAN a pour objet la conception d'un système d'information sémantique permettant à des navigateurs (marins) d'accéder aux informations maritimes pertinentes en fonction de leurs besoins. Le système d'information propose des mécanismes

de recherche d'information, une structure organisant l'accès aux informations pertinentes en fonction des tâches utilisateur et des mécanismes d'adaptation – filtrage des informations pertinentes. Les connaissances explicites de la communauté de pratiques concernée sont constituées d'un modèle de domaine qui permet d'indexer les informations, d'un modèle de tâches hiérarchiques décrivant les tâches utilisateur et organisant l'accès aux informations, d'un modèle utilisateur, d'un modèle de contexte et d'un modèle d'adaptation. La recherche d'information se fait à l'aide de requêtes utilisant principalement un moteur d'inférence. Ces requêtes sont constituées des concepts et des relations du modèle de domaine. Le système d'information est accessible sur internet au travers d'un navigateur, mais les informations ne sont réutilisables que localement. Le système d'information adaptatif est engendré dynamiquement à partir des modèles précédents.

3.2 Introduction

Les nouvelles technologies de l'information et de la communication (internet, intranet, serveurs web, ...) mettent à la disposition de tous des informations multimodales (texte, image animée ou non, vidéo, son, programmes, ...) et offrent un moyen d'accès simple et uniforme aux informations à l'aide d'un navigateur. Ce moyen d'accès offre l'avantage d'être standardisé et accessible à partir de tout poste de travail en entreprise et d'ordinateurs familiaux et ceci quel que soit le système d'exploitation. La conception des systèmes d'information s'en trouve fortement modifiée pour plusieurs raisons : d'une part ces systèmes utilisent de plus en plus ces technologies, d'autre part la quantité d'informations accessibles s'étant fortement accrue l'organisation et l'accès aux informations pertinentes pour chaque utilisateur devient primordiale⁵. Un serveur Web propose aux utilisateurs des documents hypermédia et des accès aux informations, orientés utilisateur. Les mécanismes de recherche d'informations sont fondés sur un mécanisme de navigation dans des bases d'informations au lieu d'écrire des requêtes. L'un des principaux avantages de naviguer pour les utilisateurs est qu'il est plus aisé de reconnaître les informations recherchées que de les caractériser à priori - à l'aide de requêtes (Vassileva 1996). Les hypermédias ont néanmoins quelques inconvénients : un utilisateur peut être facilement « perdu » dans l'espace d'information en naviguant de l'une à l'autre – désorientation -, et plus particulièrement lorsque la base d'information est importante (Conklin 1987). Cette désorientation nuit tout particulièrement à la compréhension du document visualisé – comme nous le verrons plus tard. Il devient donc nécessaire d'adapter l'accès à l'information en fonction des besoins utilisateur. L'adaptation - personnalisation - est l'un des enjeux importants pour la conception et l'accès à ces nouveaux systèmes d'information en ligne. Le concept d'adaptation est abordé principalement dans les deux communautés suivantes : les hypermédias adaptatifs et la modélisation utilisateur que nous allons introduire maintenant.

Le projet SWAN a été réalisé en collaboration avec le SHOM (service hydrographique et océanographique de la marine) et la société Atlantide. La conception d'un site web adaptatif quel que soit le domaine envisagé nécessite une analyse du domaine et des pratiques courantes des futurs utilisateurs afin de déterminer les points particuliers nécessitant une

⁵ Ces systèmes s'organisent souvent autour d'un portail proposant l'accès à de nombreux services autrefois séparés et bien souvent dédiés à quelques utilisateurs. De nombreux utilisateurs ayant des objectifs et caractéristiques différentes accèdent maintenant à ces nouveaux systèmes d'informations.

adaptation, les critères d'adaptation, ainsi que les méthodes à utiliser. Après une première série d'interviews auprès du SHOM, de la marine - militaire -, de la marine de commerce - d'un navire poseur de câble de France Télécom, le Léon Thévenin - et de plaisanciers, il est apparu intéressant de concevoir un système d'information intégrant les publications du SHOM (instructions nautiques, livres des feux et des signaux de brume, livres des radiocommunications, etc.) et d'adapter l'accès à ces informations en fonction des besoins utilisateur. Lors d'une deuxième série d'interviews, nous avons cherché à connaître les différents besoins utilisateur pour la préparation d'une navigation ou la navigation. Il est bien évident que des interviews libres et semi-dirigés ne permettent que d'obtenir les comportements courants et les lacunes des documents papiers par rapport à leurs besoins. A partir de ces informations, un système offrant de nouveaux objectifs s'inspirant des pratiques courantes et des besoins exprimés - non satisfaits - peut être conçu.

Ce chapitre est organisé autour de 5 paragraphes. Le premier introduit brièvement la problématique de la compréhension dans l'utilisation d'hypermédias. Le second paragraphe propose un état de l'art sur les hypermédias adaptatifs et la modélisation utilisateur. Le troisième paragraphe présente le projet SWAN dédié à la conception d'un site web adaptatif pour des navigateurs ou marins. Les deux derniers sont constitués de la conclusion et de la bibliographie.

3.3 Compréhension d'un hypermédia

L'adaptation permet aussi de résoudre l'un des problèmes importants des hypermédias : la compréhension des documents visualisés, notamment lorsque l'espace de recherche des informations est important. En science cognitive, la compréhension d'un document hypermédia est souvent caractérisée par la construction mentale d'une représentation d'un modèle de ce document. La lisibilité du document peut être définie comme l'effort mental nécessaire au processus de construction d'un modèle (Kaheneman 1973; Thüning, Hannemann et al. 1995). Afin d'accroître la lisibilité d'un document hypermédia, il est nécessaire d'assister le lecteur dans la construction de ce modèle mental, par exemple en favorisant les facteurs utiles pour le processus de construction et en diminuant les facteurs qui le perturbent. Deux facteurs principaux ont été identifiés : la cohérence et le surcoût cognitif (Thüning, Hannemann et al. 1995).

- Pour augmenter la cohérence d'un document, il faut fournir au lecteur des repères l'aidant à identifier les composants majeurs du document et sa structure globale. Pour réduire l'effort mental de compréhension, il n'est pas suffisant d'avoir une structure globale cohérente, il faut aussi communiquer cette structure au lecteur. Il est alors possible de fournir au lecteur les différents composants du document ainsi que leurs relations. La cohérence de cette structure globale peut être assurée par une organisation qui fait sens pour l'utilisateur. En d'autres termes, elle doit avoir un niveau d'abstraction suffisant pour l'utilisateur et être fondée sur les connaissances, les compétences et le savoir-faire des utilisateurs mis en œuvre dans les différents contextes d'utilisation.
- Conklin caractérise le surcoût cognitif comme l'effort supplémentaire et la concentration nécessaire pour gérer plusieurs tâches en même temps (Conklin 1987). Ce phénomène est dû aux capacités cognitives limitées du processus humain de

traitement de l'information (Kaheneman 1973). Dans un document hypermédia, ces efforts sont dus à l'orientation et la navigation. Pour éviter la désorientation, le lecteur doit non seulement connaître la structure globale du document, mais aussi ses déplacements dans cette structure. Il doit en fait savoir où il est, d'où il vient et où il peut aller. Compréhension et orientation sont en effet très fortement liées.

Dans des hypermédias ayant un grand espace d'information, le lecteur n'a besoin d'accéder qu'à une faible partie de l'espace. Par conséquent, réduire l'espace d'information aux seules informations pertinentes est un bon moyen d'éviter pour le lecteur d'être perdu et facilite la compréhension. Il est donc nécessaire de filtrer l'information pour obtenir un sous-espace de faible taille dans lequel le lecteur pourra naviguer facilement sans se perdre. Concevoir un système d'information en ligne sur internet/intranet nécessite donc d'acquérir de bonnes stratégies de filtrage associées à une structure globale de document. Cette dernière doit posséder les propriétés suivantes :

- Stabilité : un site web peut engendrer des documents électroniques à partir d'une ou plusieurs structures globales dont la durée de vie est au moins égale à celle de la session. En d'autres termes, cette structure est stable pendant la lecture d'un document, sinon le lecteur aura des difficultés de compréhension.
- Réutilisabilité : à partir du moment où le document hypermédia est engendré dynamiquement à partir d'une structure globale, des stratégies de filtrage associées, des informations disponibles et du modèle utilisateur, il est alors possible de réutiliser ces informations pour d'autres utilisateurs ou avec d'autres structures globales sans avoir à modifier de nombreux composants du système.
- Extensibilité : cette structure globale doit être aisément extensible et modifiable afin d'engendrer dynamiquement de nouveaux documents et/ou de nouvelles versions de documents sans modifier les autres parties du système.
- Compréhensibilité: son niveau d'abstraction doit être suffisant pour être interprétable de manière univoque par le lecteur. Une représentation sémantique formelle de cette structure fondée sur les objectifs et les connaissances des utilisateurs permet d'obtenir un niveau d'abstraction adéquat pour une bonne compréhension.

Une telle structure globale peut fournir dynamiquement les objectifs suivants : organisation globale de l'hypermédia, organisation de l'accès aux informations et donc fournir les différentes étapes de navigation dans l'hypermédia, vue globale de l'hypermédia, repères locaux et globaux pour faciliter la navigation et l'orientation dans le document - donc la compréhension. Cette notion de structure globale peut être instanciée de manière très différente selon le domaine d'application envisagé. Dans de nombreux systèmes hypermédia dynamiques, on utilise généralement un ou des modèles de domaine pour indexer les informations. Ce modèle pourrait servir de structure stable et interprétable par un lecteur puisqu'elle représente les concepts du domaine ainsi que leurs relations. Cependant, elle n'est pas toujours bien adaptée aux besoins utilisateur pour les raisons suivantes :

- Cette structure peut être complexe dans certains domaines d'applications.

- Les informations associées à un concept donné peuvent être beaucoup trop nombreuses et donc ne pas permettre un filtrage suffisant de l'espace d'information.

Le modèle de domaine est nécessaire pour l'indexation et le filtrage des informations mais pas suffisante. Il faut en effet concevoir des structures plus abstraites comme par exemple des modèles de tâches, des stratégies pédagogiques, des stratégies de résolution de problème, des stratégies d'aide à la décision, etc. fonction des objectifs de l'hypermédia.

3.4 Les hypermédias adaptatifs

Le principe des hypermédias adaptatifs repose sur la constatation suivante : les hypermédias traditionnels présentent les mêmes pages et les mêmes hyperliens à tous les utilisateurs. Or, ceux-ci peuvent différer les uns des autres par leurs objectifs, leurs acquis, leurs antécédents et leurs connaissances par rapport au sujet traité dans l'hypermédia. Par conséquent, ils ne seront pas intéressés par les mêmes informations et ne suivront pas les mêmes liens lors de leur navigation. L'objectif des hypermédias adaptatifs est donc d'assister les utilisateurs pour l'accès à l'information en fonction de leurs besoins (Iksal 2002). Peter Brusilovsky nous en donne la définition suivante :

- « By adaptive hypermedia systems we mean all hypertext and hypermedia systems which reflect some features of the user in the user model and apply this model to adapt various visible aspects of the system to the user » (Brusilovsky 1996).

Ces hypermédias peuvent être caractérisés d'adaptable ou d'adaptatif (Bra 1999). Dans un hypermédia adaptable, le modèle utilisateur est statique. En d'autres termes, il n'y a pas d'observateur qui mette à jour le modèle utilisateur pendant ou après une session. Dans un hypermédia adaptatif, l'observateur enregistre des informations sur le comportement utilisateur puis modifie dynamiquement le modèle utilisateur. La conjonction des deux – adaptable et adaptatif - est tout à fait possible. En effet, certaines propriétés d'un modèle utilisateur peuvent être statiques et d'autres dynamiques.

Un hypermédia est principalement composé d'une organisation (plan d'un site web) souvent représentée implicitement au travers des liens de navigation, d'un contenu et d'une présentation. On verra par la suite que le contenu et la présentation ont été longtemps confondus. L'adaptation peut être vue comme un quadruplet : contenu, modèle utilisateur, observateur et règles d'adaptation. Ce type de quadruplet a été utilisé pour une formalisation logique de l'adaptation afin de comparer différents EIAH fondés sur des hypermédia adaptatifs (Henze and Nejd 2004). Le contenu est constitué de toutes les informations qui pourront être présentées à l'utilisateur. Le modèle utilisateur est composé d'un ensemble d'informations persistantes – encore appelées « sources de connaissances » sur un utilisateur - qui caractérisent un utilisateur ou un groupe d'utilisateurs particuliers. Un tel modèle peut contenir des caractéristiques sur les connaissances, les préférences, les objectifs, les centres d'intérêts, ... d'un utilisateur. L'observateur enregistre certaines informations sur le comportement de l'utilisateur et modifie en conséquence son modèle. Le rôle de l'adaptation est de fournir le contenu, la navigation – fondée sur l'organisation - et la présentation qui sont pertinents en fonction des besoins utilisateur. Pour cela, les règles d'adaptation peuvent utiliser la description des contenus, des organisations et des présentations et le modèle utilisateur pour décider de leurs pertinences et ensuite appliquer une méthode d'adaptation.

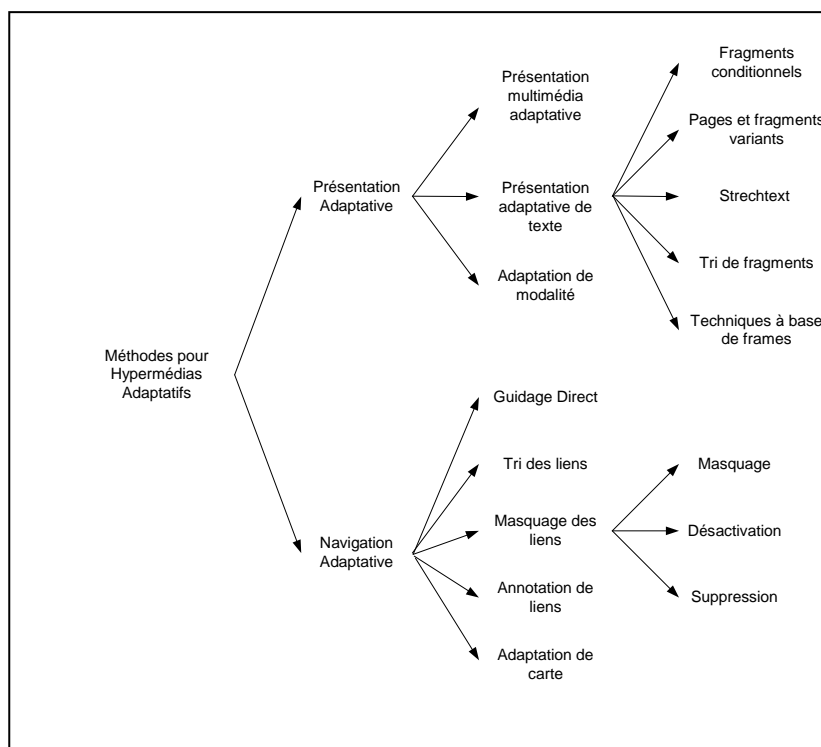


Figure 20 : Méthodes d'adaptation

Dans la cartographie des techniques d'adaptation présentée dans la figure 20, nous retrouvons les deux catégories d'adaptation les méthodes de présentation adaptative qui traitent le contenu des hypermédias et les méthodes de navigation adaptative qui concernent plus particulièrement les liens (Bra 1999; Brusilovsky 2001)⁶. Le but de la présentation adaptative est d'adapter le contenu d'une page accédée en fonction du modèle utilisateur. La navigation adaptative a pour but d'aider les utilisateurs à trouver leur chemin dans l'espace d'information en adaptant la présentation des liens et/ou en supprimant certains liens en fonction du modèle utilisateur. Dans la majorité des cas, c'est le nœud destination d'un lien qui est évalué pour déterminer la pertinence d'un lien et non pas le lien lui-même.

Nous allons maintenant présenter chacune des techniques de cette cartographie en commençant par celles qui adaptent la présentation, puis nous enchaînerons sur celles qui adaptent la navigation.

3.4.1 Les méthodes de présentation adaptative

Selon Peter Brusilovsky (Brusilovsky 1996), il existe cinq méthodes de présentation adaptative :

⁶ Le terme de présentation adaptative n'est pas très juste puisqu'il s'agit principalement d'adaptation de contenu. Comme les premiers hypermédias adaptatifs étaient constitués de pages html statiques dont le contenu était prédéterminé, l'adaptation consistait principalement à masquer, sélectionner, trier ou annoter des contenus. Elle a donc été nommée présentation adaptative, sans doute parce que son traitement se faisait au niveau de la présentation. Mais elle a masqué la possibilité d'adapter réellement la présentation pendant quelques années.

- L'explication additionnelle : les seules parties du document qui sont présentées à l'utilisateur sont celles qui correspondent par exemple à son niveau de connaissance.
- L'explication pré-requise : le système évalue les pré-requis nécessaires à la compréhension de la page présentée à l'utilisateur. Si ce dernier ne possède pas les connaissances suffisantes, le système intègre des informations supplémentaires sur les concepts méconnus (ou inconnus) dans la page.
- L'explication comparative : l'idée des explications comparatives est de présenter à l'utilisateur un concept inconnu en le comparant à un concept similaire qui lui est familier - connu. Des explications insistant sur les similarités et les différences entre les deux concepts lui sont proposées.
- Les explications variantes : le système possède différentes versions - fragments – pour un même concept. Il sélectionne celui qui correspond le plus au modèle de l'utilisateur.
- Le tri : les différents fragments d'information sont triés en fonction de leur pertinence pour l'utilisateur.

L'implémentation de ces différentes méthodes se fait au travers des techniques suivantes :

- Le texte conditionnel : l'affichage d'un fragment de texte n'est réalisé que lorsque la condition associée est satisfaite. Il s'agit d'une comparaison entre le contenu – texte – et la modèle utilisateur. Cette technique a été utilisée dans ITEM/IP (Brusilovsky, Pesin et al. 1993), Lisp-Critic (Fischer 1990), C-book (Kay and Kummerfeld 1994) et AHA (Bra and Calvi 1998).
- Le stretchtext : certains mots clés d'un texte peuvent être remplacés par une description plus longue si la connaissance de l'utilisateur nécessite ce développement. L'hypermédia propose des liens activables pour certains mots clés du texte. L'activation de ces liens ajoute un texte supplémentaire et permet donc à l'utilisateur de disposer d'un texte plus complet. Cette technique est utilisée par exemple dans MetaDoc (Boyle and Encarnacion 1994) et dans KN-AHS (Kobsa, Müller et al. 1994).
- La réorganisation de fragments : le système peut réorganiser les fragments présentés à l'utilisateur en fonction de ses préférences ou de ses objectifs. Par exemple, certains utilisateurs préféreront la définition avant l'exemple, et d'autres l'inverse. Le système peut aussi trier les fragments en fonction de leur pertinence pour l'utilisateur, du plus pertinent au moins pertinent.
- Les pages et fragments variants : le système dispose de plusieurs versions d'une même page ou d'un même fragment chacun associé à un stéréotype d'utilisateur. Au moment d'afficher une page, l'hypermédia sélectionne celle dont le stéréotype est associé à l'utilisateur.
- Les techniques à base de frames : l'ensemble de l'information sur un concept est représenté sous la forme d'un frame. Les slots du frame contiennent des fragments variants, des liens vers d'autres frames, ou encore des exemples. Dans le projet EPIAIM (Rosis, Carolis et al. 1993), les règles permettent de choisir des schémas qui

sont en fait des ensembles ordonnés de slots. Le projet PUSH (Höök, Karlgren et al. 1996) utilise une technique peu répandue qui combine le stretchtext et les frames.

- L'adaptation de modalité : le système dispose de plusieurs présentations d'une information, chacune correspondant à un média différent. En fonction des préférences de l'utilisateur et du contexte de consultation (support, logiciels disponibles), le système propose l'information sur l'un ou l'autre des médias.

Pour la présentation adaptative, les règles d'adaptation comparent la description du contenu au modèle utilisateur pour les contenus dynamiques – résultat d'un processus de recherche d'information. Pour les contenus statiques - prédéterminés -, les règles d'adaptation n'utilisent que le modèle utilisateur pour évaluer la pertinence du contenu. Il n'y a pas de description du contenu. Le concepteur le connaît et associe la règle d'adaptation à celui-ci.

3.4.2 Les méthodes de navigation adaptative

Peter Brusilovsky définit la navigation adaptative autour de cinq méthodes qui peuvent être aussi considérées comme des buts à atteindre par des techniques de navigation adaptative (Brusilovsky 1996).

- Le guidage global : le but du guidage global est de permettre à l'utilisateur de prendre le chemin le plus court pour accéder à l'information qu'il recherche. Il est nécessaire que l'utilisateur ait une idée générale de l'information recherchée et que l'organisation serve de support à la navigation.
- Le guidage local : le but du guidage local est d'aider l'utilisateur à passer d'une page à l'autre en lui suggérant le lien le plus intéressant à suivre à partir du nœud courant.
- Le repère de navigation global : le repère de navigation global a pour objectif d'aider l'utilisateur à comprendre la structure globale de l'hypermédia - ou organisation - ainsi que sa position dans celui-ci. Cette organisation se doit de représenter d'une manière ou d'une autre les objectifs globaux de l'utilisateur.
- Le repère de navigation local : le repère de navigation local a pour objectif l'orientation de l'utilisateur dans l'hypermédia. Il lui indique où il est, d'où il vient et aussi où il peut aller.

Ces techniques de navigation adaptative utilisent la structure globale de l'hypermédia et ont aussi pour objet de communiquer cette dernière au lecteur, ainsi que ses déplacements au sein de cette structure en les adaptant aux besoins utilisateur. Elles permettront donc de faciliter la compréhension de l'hypermédia engendré.

Ces méthodes d'adaptation peuvent être réalisées à partir des techniques suivantes :

- Le guidage direct : le système détermine à partir du modèle utilisateur le meilleur "prochain" nœud à visiter. Le lien vers ce nœud est représenté par un bouton "suivant" ou "continuer" (Brusilovsky and Pesin 1994).

- Le tri : il s'agit de guider l'utilisateur en lui proposant une liste de liens triés du plus pertinent au moins pertinent pour l'utilisateur. Cette technique est très utilisée dans le domaine de la recherche d'information, par exemple dans l'outil HYPERFLEX (Kaplan, Fenwick et al. 1993).
- Le masquage : c'est la technique qui est la plus transparente pour l'utilisateur, il s'agit de masquer les liens non pertinents afin de réduire la surcharge cognitive. La pertinence des liens est évaluée à partir du modèle utilisateur. Le terme masquage peut être interprété de différentes façons, Paul De Bra et Licia Calvi, dans leur système AHA (Bra and Calvi 1998), affinent la notion de masquage de la façon suivante :
 - Le masquage de liens : les liens non pertinents ne sont plus visibles. Ils sont présentés comme du texte "normal".
 - La désactivation de liens : le lien est toujours visible, mais l'utilisateur ne peut pas l'activer.
 - La suppression de liens : cette technique consiste à supprimer du document les liens non pertinents. Il s'agit d'une technique très intéressante dans le cas de listes de liens, mais elle est inutilisable lorsque le lien se situe au sein d'un texte.
- L'annotation : cette technique consiste à présenter les liens différemment en fonction de la pertinence du nœud destination – son contenu - pour l'utilisateur. Cette technique permet de classer des liens dans différentes catégories. Pour chaque classe, une annotation particulière est associée au lien – icônes, couleurs, etc.

En règle générale, la détermination de la pertinence d'un lien se fait de deux manières différentes : soit en évaluant le lien à partir de sa description ou de son type lorsqu'ils sont typés, soit en évaluant le contenu du nœud destination. La combinaison des deux est bien évidemment possible. Mais, dans la plupart des cas, l'évaluation se fait par celle du contenu du nœud destination. Dès que la technique de navigation adaptative supprime ou ne donne plus accès à des liens, il y a bien évidemment une modification du contenu de l'hypermédia, donc une adaptation de contenu. La présentation adaptative et la navigation adaptative sont donc liées dans certains cas.

En conclusion, les règles d'adaptation ont pour objet de comparer des contenus ou leur description à un modèle utilisateur. Si le contenu est statique, c'est-à-dire prédéterminé, il suffit de ne prendre en compte que le modèle utilisateur dans les règles d'adaptation. Il n'y a pas de description du contenu. Le concepteur le connaît et associe la règle d'adaptation à celui-ci. Par contre, si les contenus sont le résultat d'un processus de recherche d'information, alors l'adaptation peut être vue comme un raffinement de la recherche d'information. Ce point de vue peut s'appliquer pour la présentation adaptative et la navigation adaptative qui affectent respectivement le contenu et l'organisation dont la pertinence des liens est presque toujours évaluée à l'aide du contenu du nœud destination. Le contenu et/ou sa description, les règles d'adaptation et le modèle utilisateur sont fortement dépendants. En effet, les règles d'adaptation utilisent les propriétés du contenu et du modèle utilisateur dédiées à l'adaptation et/ou à un type d'adaptation. Nous allons maintenant présenter les modèles utilisateurs.

3.4.3 La modélisation utilisateur

Adapter, personnaliser un document ou une application pour un utilisateur particulier nécessite de disposer d'informations sur ce dernier permettant d'évaluer la pertinence des objets disponibles (contenu multimédia, organisation - liens de navigation - et présentation) ou d'aider le système à faire des choix. Le modèle utilisateur est « une source de connaissances, une base de données sur un utilisateur » (McTear 1993). Plus précisément, il s'agit d'un ensemble d'informations persistantes qui caractérisent un utilisateur ou un groupe d'utilisateurs particuliers. Un tel modèle peut contenir des caractéristiques sur les connaissances, les préférences, les objectifs, les centres d'intérêts, etc. d'un utilisateur. Les caractéristiques d'un modèle utilisateur peuvent être analysées individuellement ou en sous-groupes selon les quatre dimensions principales suivantes : les modes d'acquisition, le type du modèles, les catégories de caractéristiques, les propriétés de ces caractéristiques (Iksal 2002). Nous allons maintenant les détailler toutes.

3.4.3.1 L'acquisition des caractéristiques

On dispose de deux manières d'obtenir de l'information sur les utilisateurs : l'acquisition explicite, une source externe à l'outil est utilisée pour créer et/ou compléter le modèle et l'acquisition implicite, le système infère l'information à partir des connaissances disponibles sur l'utilisateur. Parmi les méthodes explicites les plus souvent utilisées on peut citer :

- L'observation directe : il s'agit de la méthode la plus précise. Elle permet d'identifier des classes d'utilisateurs ainsi que les tâches de ces derniers. De plus, on peut identifier des facteurs critiques, comme la pression sociale, qui ont des effets néfastes sur l'utilisation du système. Malheureusement, il s'agit d'une méthode très coûteuse qui nécessite des personnes qualifiées derrière chacun des individus observés.
- Les interviews : cette technique permet d'obtenir un autre type d'information, l'expérience, les opinions, les motivations comportementales mais surtout les avis sur les outils existants. Ils sont plus courts et moins coûteux que la technique d'observation, néanmoins, ils nécessitent aussi du personnel qualifié.
- Les questionnaires : « comment obtenir à moindre coût un maximum d'informations » (Petrelli, Angeli et al. 1999) ? Les résultats obtenus permettent des études statistiques et des généralisations plus fortes que les interviews. Les questionnaires peuvent être collectés par des personnes non expérimentées. Ils permettent d'avoir à la fois un aperçu de la situation et des points d'information plus précis.

La situation idéale, qui est d'utiliser à la fois ces trois méthodes, a déjà été pratiquée - (Vassileva 1996) - mais est rarement justifiée. Le plus souvent une seule de ces techniques suffit à rassembler les informations nécessaires pour le modèle utilisateur d'une application.

Une des méthodes implicites la plus utilisée est l'apprentissage. Il est alors possible d'obtenir de l'information sur un ou des utilisateurs. En effet, lorsqu'il existe déjà un système, il est intéressant et très utile d'y ajouter un algorithme d'apprentissage pour obtenir des informations essentiellement comportementales sur l'utilisateur (Pohl 1996).

3.4.3.2 Les types de modèles

Un modèle utilisateur peut être constitué à partir des différents types de modèles que nous allons présenter ici. En d'autres termes, il est possible de les combiner pour la création d'un unique modèle utilisateur :

- Le modèle individuel : ce modèle regroupe les caractéristiques propres à un individu.
- Le stéréotype : il s'agit d'un modèle générique qui correspond à un condensé des caractéristiques les plus représentatives d'un groupe - ou classe - d'individus, que nous pouvons qualifier de valeurs par défaut. Le stéréotype peut être utilisé tel quel, ou dans une optique d'individualisation de modèle (Rich 1989).
- Le modèle de recouvrement : il ne peut exister que si le système possède un modèle de domaine car le modèle de recouvrement associe pour un utilisateur particulier, une valeur à chaque concept du modèle de domaine. Chaque valeur correspond à une estimation du niveau de connaissance de l'utilisateur pour le concept. Le modèle de recouvrement est facile à mettre à jour mais souvent difficile à initialiser surtout au niveau de la mesure du niveau de connaissance. Les actions de l'utilisateur sont analysées pour augmenter ou réduire le niveau de connaissance des concepts du domaine (Brusilovsky, Eklund et al. 1998).
- Le modèle partagé : il s'agit d'un modèle réutilisable dans plusieurs applications. En effet, de nombreuses caractéristiques d'un utilisateur sont utilisées systématiquement par toutes les applications adaptatives. L'idée du modèle partagé est de disposer d'une partie commune et de parties spécifiques à chacune des applications ou tâches à réaliser. Cette approche permet d'une part de partager des informations entre applications et d'autre part d'obtenir plus facilement une participation des utilisateurs qui n'ont pas à redéfinir un modèle pour chacune des applications. Kobsa et Wahlster (Kobsa and Wahlster 1989) proposent de disposer d'un modèle central et de plusieurs sous-modèles disjoints.

3.4.3.3 Les caractéristiques de l'utilisateur

L'utilisateur peut être modélisé par de très nombreuses caractéristiques différentes. D'une certaine manière, il n'y a pas vraiment de limite au nombre et au type de celles-ci. Seuls les objectifs de l'adaptation pour l'hypermédia considéré déterminent les caractéristiques pertinentes. Il existe malgré tout un certain nombre de caractéristiques couramment utilisées que nous allons présenter ici. Celles-ci sont : la connaissance, l'expérience, les compétences, les préférences et les objectifs.

3.4.3.3.1 La connaissance

La connaissance de l'utilisateur apparaît comme la caractéristique la plus utilisée dans les hypermédias adaptatifs. La façon la plus simple de gérer les connaissances est de mémoriser ce que l'utilisateur connaît ou ne connaît pas, comme par exemple un concept provenant d'un modèle de domaine. Cette information peut être obtenue explicitement par l'utilisateur ou implicitement par le système sur la base de règles d'inférences. La connaissance est un

paramètre dynamique pour un utilisateur particulier, en effet, l'utilisateur apprend de nouvelles informations chaque jour. Les modèles de connaissances posent le problème de la nature de la connaissance (Kobsa and Pohl 1995). En effet, ce qui est modélisé peut être : ce que l'utilisateur connaît effectivement, ce que l'utilisateur croit connaître, ce que le système considère comme connu par l'utilisateur. Il est donc nécessaire de disposer de méthodes d'évaluation de la connaissance permettant d'affirmer qu'une croyance (du système ou de l'utilisateur) est en fait une connaissance effective. On dispose de deux approches pour modéliser la connaissance ; elles ne sont pas exclusives. Les stéréotypes permettent de modéliser rapidement la connaissance par défaut que possède un individu appartenant à une certaine classe d'utilisateurs. Parfois les stéréotypes sont hiérarchisés, en analysant un modèle utilisateur individuel, il est alors possible de classer un individu sous un ou plusieurs stéréotypes à la suite d'un certain nombre d'interactions. Un système adaptatif peut se contenter des stéréotypes, mais il peut aussi les utiliser pour initialiser des modèles de recouvrement. Il s'agit dans ce cas de modéliser la connaissance d'un individu et non celle d'un groupe. Ce type de modèle est puissant et flexible. En effet, il est possible de faire évoluer assez facilement le niveau de connaissance de l'utilisateur pour chacun des concepts. Dans certains cas, le stéréotype et le modèle de recouvrement ont la même structure, seulement le stéréotype donne le niveau minimal de connaissance qu'une classe d'utilisateurs doit avoir sur un concept, alors que le modèle de recouvrement « devrait représenter exactement » le niveau de connaissance d'un individu.

3.4.3.3.2 L'expérience, les compétences

Deux caractéristiques sont similaires à la connaissance définie dans le paragraphe précédent mais diffèrent de par la nature même de l'information qu'ils représentent :

- L'expérience : l'expérience de l'utilisateur représente son savoir-faire, la familiarité et l'aisance qu'il possède avec le type de système qui lui est présenté. Pour un hypermédia adaptatif, l'expérience peut représenter les capacités de l'utilisateur à comprendre la structure hypertextuelle ainsi que celles de la navigation dans un hypermédia.
- Les compétences : les compétences possédées par l'utilisateur correspondent aux connaissances qui ne relèvent ni du domaine, ni de l'expérience mais qui sont néanmoins considérées comme pertinentes dans le fonctionnement du système. Prenons l'exemple d'un étudiant en économie qui consulte un hypermédia sur l'Egypte Ancienne. Les connaissances évoquées dans le paragraphe précédent concernent l'Egypte Ancienne, alors que l'expérience porte sur les systèmes hypermédia et que les compétences peuvent regrouper ses connaissances en économie, en politique, etc.

3.4.3.3.3 Les préférences

Pour différentes raisons, l'utilisateur préfère une interface plutôt qu'une autre, une technique d'adaptation particulière (masquage, annotation, etc.), ou encore, un auteur particulier, un type de littérature (roman, science fiction, etc.). Les préférences de l'utilisateur peuvent porter sur une généralité et/ou un point particulier. Par exemple, un voyageur exige le côté fenêtre systématiquement, sauf pour l'avion, où il demande le couloir. Contrairement aux autres caractéristiques, les préférences ne peuvent être déduites par le système. En effet,

chaque utilisateur possède ses propres préférences, et c'est à lui de les renseigner. Les préférences sont utilisées par le système soit à des fins d'adaptation, de sélection de stéréotypes, ou encore pour inférer des hypothèses sur l'utilisateur (Kobsa 1993). Mais elles permettent aussi de combiner différents modèles utilisateur individuels pour en extraire des modèles de groupes qui serviront de base lors de l'arrivée de nouveaux membres dans le groupe (Brusilovsky 1996).

3.4.3.3.4 Les objectifs

Dans les systèmes d'information en ligne, les utilisateurs ont l'habitude de n'accéder qu'à une partie de l'espace d'information - ensemble de fragments - qui dépend principalement de leurs objectifs (Brusilovsky 1996; Vassileva 1996). L'objectif est un état que l'utilisateur espère atteindre, et les plans décrivent les étapes pour y arriver. Les modèles de tâches sont très appropriés pour modéliser ces plans et par conséquent les objectifs des utilisateurs. Souvent ces démarches sont typiques d'une catégorie d'utilisateurs, c'est pourquoi on associe un modèle de tâches statique à un stéréotype.

Ces informations peuvent être données explicitement par l'utilisateur, ou alors apprises par le système, grâce à des outils de reconnaissance de plans. Une difficulté avec l'apprentissage, vient du fait que l'objectif d'un utilisateur peut changer et que par conséquent le plan aussi, dans ce cas on peut proposer à l'utilisateur un éditeur de modèles de tâches composé d'un catalogue de tâches qu'il est possible d'ordonnancer, afin qu'il donne explicitement une séquence de tâches qui correspond à sa démarche.

Il n'est envisageable d'utiliser une telle méthode que lorsque les objectifs utilisateur ont une certaine stabilité. En effet, la tâche réelle d'un utilisateur est rarement identique à celle prescrite par le système. Le modèle n'est qu'un guide rendant compte d'une partie de celle-ci, mais qui doit être suffisamment proche pour être utile. L'objectif est une caractéristique qui peut changer souvent au fil des sessions, mais aussi au sein d'une même session. On peut donc considérer deux types d'objectifs pour un utilisateur : les objectifs généraux qui sont stables pour une session donnée et les objectifs locaux que l'utilisateur peut décider de modifier en cours de session. Les objectifs généraux peuvent être considérés comme des orientations dans l'utilisation du système. La modélisation de tâches et plus particulièrement l'utilisation de modèles de tâches hiérarchiques pose le problème du niveau de granularité de la modélisation. Si les tâches sont trop générales, elles ne sont pas d'un grand recours pour l'utilisateur. Et si elles sont trop spécialisées, c'est-à-dire d'un niveau de granularité trop fin, elles deviennent une gêne pour l'utilisateur. Il risque en effet d'y avoir un conflit permanent entre tâche réelle et tâche prescrite.

3.4.3.4 Les propriétés des caractéristiques utilisateurs

Les caractéristiques modélisées possèdent des propriétés soit en raison de leur acquisition, soit en raison de leur utilisation. Nous allons présenter ici les principales propriétés avec des questions qu'il est nécessaire de se poser :

- Quel est le sujet modélisé ? On peut modéliser un individu ou une classe d'individus (par exemple en utilisant des stéréotypes), l'utilisateur direct du système ou le bénéficiaire du résultat obtenu. Dans HAM-ANS (Hahn and Wahlster 1983), la

réservation de la chambre d'hôtel n'est pas nécessairement faite par l'intéressé, mais peut être demandée par sa secrétaire. Dans ce cas, certaines caractéristiques concerneront la secrétaire et d'autres concerneront le client de l'hôtel.

- La caractéristique est-elle stable ? Les caractéristiques considérées comme stables sont celles qui ne peuvent pas être remises en question ou qui sont rarement modifiées. Par exemple, le nom d'un utilisateur, son activité professionnelle, sa localisation géographique, etc.
- Quelle est sa durée de vie ? Un modèle utilisateur peut contenir des informations dont la validité se limite à la session courante, ou des informations qu'il est nécessaire de conserver pour une durée plus longue.
- Dans quel cadre est utilisée la caractéristique ? Toutes les informations contenues dans un modèle utilisateur ne sont pas toujours utilisées. Certaines caractéristiques sont utilisées par le système pour des tâches particulières : recherche d'information, identification, personnalisation, etc.
- Peut-elle évoluer ? En d'autres termes, est-elle statique ou dynamique ? Certaines caractéristiques sont amenées à évoluer dans le temps. Souvent dans les systèmes tuteur la connaissance, que possède l'utilisateur, est amenée à s'enrichir au fil des utilisations. Par exemple, l'utilisateur peut passer de novice à expert sur certains concepts du domaine.
- Qui renseigne les caractéristiques ? Certaines sont renseignées par l'utilisateur parce qu'il est le seul à les connaître comme les préférences, mais d'autres peuvent être déduites par le système, comme par exemple la connaissance qui peut être déduite de l'activité professionnelle de l'utilisateur. L'évolution de certaines caractéristiques de l'utilisateur peut avoir un rythme très différent : cours terme, moyen terme ou long terme.

En conclusion, Nous avons analysé un certain nombre de caractéristiques des modèles utilisateurs. Néanmoins, cette analyse n'est pas exhaustive. Il en existe bien d'autres, mais ces caractéristiques représentent celles qui sont les plus utilisées. Le développement croissant des hypermédias adaptatifs entraîne un fort accroissement des caractéristiques utilisées dans ces modèles. Cette étude de l'adaptation et des modèles utilisateurs n'a pas pris en compte les notions de contexte qui sont de plus en plus présentes dans ce type de systèmes. Le contexte prend généralement en compte des caractéristiques non dépendantes de l'utilisateur qui jouent un rôle prépondérant dans l'adaptation. On peut citer parmi celles qui sont les plus utilisées : les informations de géo-localisation. Dans le cadre du projet SWAN, ces informations sur le contexte sont nécessaires au traitement de l'adaptation.

3.5 Le projet SWAN

L'un des objectifs principaux du projet est de concevoir un système d'information qui intègre les publications du SHOM (instructions nautiques, livres des feux et des signaux de brume, livres des radiocommunications, etc.) et qui adapte l'accès à ces informations en fonction des besoins utilisateur. A partir des interviews et de l'analyse des pratiques, on a pu

montrer que les marins avaient au moins cinq objectifs communs pour préparer un transit ou naviguer en mer : recherche et création de routes, information sur une route, port/mouillage, atterrissage et recherche d'informations générales. Les différents objectifs ont pour objet :

- Recherche et création de route : cet objectif a pour objet de fournir des routes préétablies entre deux ports / mouillages ou la création de nouvelles routes. En effet, un navire donné suit toujours la ou les mêmes routes pour aller d'un port/mouillage à un autre. Il est donc intéressant de mémoriser ces routes afin de pouvoir y avoir accès pour préparer un transit ou pour naviguer. Ces routes se composent au moins des coordonnées des différents « waypoints » et des caps compas correspondants. Dans le cadre de cet objectif, les routes et les segments de routes constituent des catégories pertinentes d'informations. Les routes pertinentes sont fonction du navire et de la réglementation correspondante, des conditions météo, des horaires de marée et parfois des préférences du navigateur.
- Recherche d'information sur une route : les instructions nautiques et livres des feux et signaux de brumes sont structurés en zones géographiques organisées dans l'ordre du serpent. Pour une route donnée, il est nécessaire de parcourir un ou plusieurs livres et de récolter des informations non contiguës pour obtenir les données relatives à une route. Les navigateurs ont pour habitude de réaliser des fiches de traversée en photocopiant, par exemple, les informations pertinentes sur une route et en les ordonnant en fonction de celle-ci. Les informations pertinentes dépendent de la route choisie, des conditions de navigation (météo, jour ou nuit, ...), du navire et de la catégorie d'information demandée. Les informations sur une route correspondent à un sous-espace du domaine. Les informations sont aussi fonction des zones géographiques des différents segments de la route choisie. Il s'agit dans cet objectif de rechercher l'information pour naviguer et de la présenter au navigateur dans un ordre compatible avec la route choisie et les conditions de navigation.
- Atterrissage : il s'agit ici de fournir les informations pertinentes pour un atterrissage⁷. Celles-ci sont principalement fonction de la zone géographique et font référence à des catégories d'informations particulières. En effet, il y a une ou plusieurs zones d'atterrissage par instruction nautique qui sont regroupées dans un chapitre particulier contenant les informations relatives à un accès du large par tous les navires. Seules les conditions de navigation ont un impact sur la pertinence des informations.
- Port/Mouillage : cet objectif permet de rechercher un port à partir d'un ensemble de critères. Les navires en transit ont parfois besoin de dévier leur route afin d'assurer du ravitaillement, des réparations, des soins, du chargement de marchandises ou matériels. Il est donc nécessaire de trouver un port capable d'accueillir le navire, mais aussi adéquat pour les opérations demandées qui dépendent souvent des infrastructures portuaires ou de la zone environnante - héliport, aéroport, autoroutes, ... Pour cela il est nécessaire de pouvoir sélectionner les ports les plus proches satisfaisants les critères requis par ces opérations. Il fournit aussi toutes les informations relatives à la

⁷ L'atterrissage est le moment où la navigateur aperçoit la côte en venant du large, passage de la navigation hauturière à la navigation côtière.

navigation dans un port pour sortir ou entrer dans celui-ci. La partie recherche de port ne sera pas implantée dans cette première version. Les informations pertinentes sont fonction du navire, de la réglementation et des conditions de navigation.

- Recherche d'information : cet objectif fournit aux navigateurs la possibilité de rechercher des informations, dans les instructions nautiques, relatives à la préparation d'une navigation telles que réglementation, statistiques de météo, nature des fonds, ...

Comme nous allons le voir maintenant, ces objectifs ont été représentés à l'aide d'un modèle de tâches hiérarchiques.

3.5.1 Modélisation des objectifs et de l'organisation

Dans les systèmes d'information en ligne, les utilisateurs n'accèdent qu'à une partie de l'espace d'information - ensemble de fragments - qui dépend principalement de leurs objectifs (Waern 1986; Tyler and Treu 1989; Grunst 1993; Kaplan, Fenwick et al. 1993; Encarnação 1995; Brusilovsky 1996; Höök, Karlgren et al. 1996; Vassileva 1996). La modélisation de ces objectifs a été notamment fondée sur des modèles de tâches hiérarchiques. Ces derniers ont souvent été utilisés en interaction homme/machine pour l'évaluation et l'interprétation (Tricot and Nanard 1997) ainsi que la conception d'hypermédia (Sebillotte and Scapin 1994; Sebillotte 1995; Fraissé 1997; Garlatti 1997; Tricot and Nanard 1997). Il n'est envisageable d'utiliser une telle méthode que lorsque les objectifs utilisateur ont une certaine stabilité. En effet, la tâche réelle d'un utilisateur est rarement identique à celle proposée par un système. Le modèle n'est qu'un guide rendant compte d'une partie de celle-ci, mais qui doit être suffisamment proche pour être utile. Pour chaque objectif utilisateur, il est possible de déterminer les informations pertinentes, puis le ou les mécanismes de filtrage associés - adaptation.

Dans ce type de systèmes d'information, les modèles de tâches auront plusieurs rôles : d'une part ils vont servir de modèle pour l'organisation du système d'information et plus particulièrement ils organisent l'accès à l'information, mais c'est aussi une partie du modèle utilisateur en tant que stéréotype.

Une analyse des différents objectifs a montré qu'il était tout à fait naturel de les représenter par un modèle de tâches hiérarchiques en raison de la structure de ces tâches mais aussi en raison de leur stabilité pour chaque utilisateur et entre les utilisateurs (Garlatti, Iksal et al. 1999). Dans ce modèle, il existe deux types de tâches : abstraites et atomiques. La structure hiérarchique du modèle de tâche est fondée sur une relation de composition entre tâches. Cette structure forme un graphe orienté sans circuit, dont les feuilles sont des tâches atomiques et les nœuds sont des tâches abstraites - chacune ayant ses propres caractéristiques et fonctionnalités. Une tâche abstraite se décompose récursivement en sous-tâches qui peuvent être elles-mêmes des tâches abstraites ou des tâches atomiques. Une structure de contrôle détermine l'ordre d'exécution des sous-tâches. Cette structure de contrôle utilise des opérateurs standards tels que la séquence (et), la sélection (xor) et une sélection particulière (ou) qui n'autorise qu'une seule tâche en attente et uniquement pour les tâches de plus haut niveau, c'est-à-dire les tâches les plus abstraites - les objectifs. Par opposition, une tâche atomique est assimilée à une feuille de l'arbre, elle n'est pas décomposable en sous-tâche.

Elle n'a aucun contrôle sur la navigation mais sert de point d'entrée pour les sous-espaces d'information ainsi que d'interface avec l'utilisateur.

Le rôle principal d'une tâche abstraite et de ses sous-tâches est d'organiser l'accès à l'information et aux interactions au travers des tâches atomiques. Comme tout modèle d'organisation dans ce type de système d'information elle fournit une aide à la navigation dans l'hypermédia. Il permet d'établir des guides et orientations locaux et globaux. En d'autres termes, elle fournit les différentes étapes de navigation pour aider l'utilisateur à réaliser son objectif et structure l'accès à l'information de façon à le rendre plus simple et plus intuitif. Par conséquent, la tâche abstraite permet aussi d'augmenter la cohérence du document hypermédia et de réduire la surcharge cognitive. Une tâche atomique remplit principalement les deux rôles suivants : interaction homme/machine et recherche d'information. Ces tâches de recherche d'information ont notamment pour objet de lier la sélection et l'organisation afin de proposer aux utilisateurs les informations pertinentes relativement à la tâche en cours. L'objet de l'adaptation est principalement de fournir à l'utilisateur les informations pertinentes pour préparer son transit ou pendant son transit. Il s'agit donc de filtrer l'information en fonction de ces besoins qui sont notamment caractérisés par la tâche courante. C'est donc dans ces tâches de recherche d'information que l'adaptation sera réalisée. Chaque tâche de recherche d'information détermine une vue de l'espace d'information contenant les plus pertinentes et ainsi offre un sous-espace d'information que l'utilisateur peut parcourir - dans lequel il est capable de se repérer. Enfin, elle détermine les méthodes d'adaptation et leurs paramètres.

Dans la suite de ce paragraphe, nous ne présenterons que deux objectifs : recherche et création de route et Information sur une route.

3.5.1.1 Objectifs de recherche et de création de routes et d'information sur une route

Nous allons ici présenter une tâche qui réunit ces deux objectifs en une seule tâche. L'objectif de recherche et de création de routes a pour objet de fournir des routes préétablies entre deux ports/mouillages ou la création de nouvelles routes. En effet, un navire donné suit toujours la ou les mêmes routes pour aller d'un port/mouillage à un autre. Il est donc intéressant de mémoriser ces routes afin de pouvoir y avoir accès pour préparer un transit ou pour naviguer. Dans le cadre de cet objectif, les routes et les segments de routes constituent les catégories d'informations recherchées. Les routes pertinentes sont fonction du navire et de la réglementation correspondante, des conditions météo, des horaires de marée et parfois des préférences du navigateur.

Pour la recherche d'information sur une route, Il s'agit de rechercher des informations pertinentes pour naviguer et de les présenter au navigateur dans un ordre compatible avec la route choisie et les conditions de navigation. Elles dépendent de la route choisie, des conditions de navigation (météo, jour ou nuit, ...), du navire et de la catégorie d'information demandée – sous-catégorie du modèle de domaine. Les informations sont aussi fonction des zones géographiques des différents segments de la route choisie. Cet objectif a donc pour objet d'aider l'utilisateur à parcourir l'espace d'information lié à sa route. Les informations fournies proviendront des Instructions Nautiques (IN), des Instructions Nautiques Plaisance, des livres des feux et signaux de brumes.

Si aucune route n'est disponible dans le système pour le transit envisagé, un objectif de création de route est proposé. Cet objectif se compose de deux sous-objectifs : l'objectif Recherche et l'objectif Création (cf. Figure 21). Comme pour tous les objectifs, il y a une étape de présentation et de terminaison à des fins de démonstration pour le prototype conçu. Avant de fournir les informations sur une route, il est nécessaire de choisir ou de créer une route. Cet objectif se compose donc des deux parties principales : sélection d'une route et recherche d'information sur celle-ci. La tâche est organisée de la manière suivante (cf. fig. 21) :

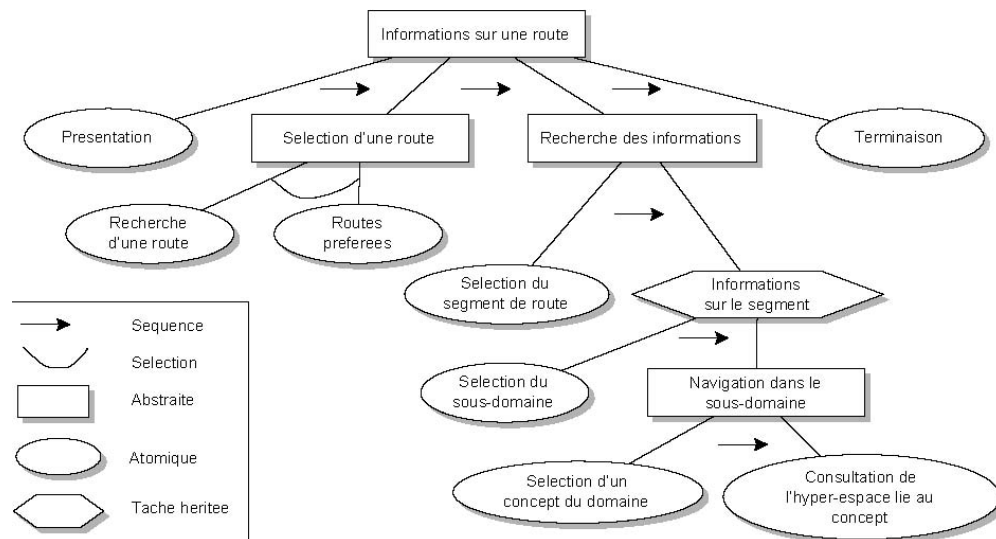


Figure 21 : Décomposition d'un objectif recherche/création de routes

Après sélection d'une route, l'utilisateur peut accéder à la tâche de recherche d'information sur une route. Il devra avant tout choisir un segment de route, ce qui peut être réalisé automatiquement à partir de la position estimée du navire et/ou du GPS. Après avoir choisi une catégorie d'information – sous-catégorie du modèle de domaine -, il pourra parcourir l'espace d'information retreint proposé par le système. On voit bien qu'il a deux étapes dans le filtrage de l'information : la première qui est conceptuelle et l'autre qui est spatiale.

Une route se compose de plusieurs segments. Chaque segment est typé, c'est-à-dire qu'il peut appartenir aux types suivants : entrée/sortie port, atterrissage, navigation côtière, navigation hauturière – ce sont des « types de navigation » décrits dans les instructions nautiques. Pour chaque type de segment, il y a deux grandes catégories d'informations qui sont les aides à la navigation - balisage, amers, alignement, feux et signaux de brumes - qui servent de repères - physiques - pendant la navigation et le contenu des instructions nautiques. L'accès et la structuration ainsi que les méthodes de présentation et d'adaptation de ces deux catégories d'information n'étant pas les mêmes, deux sous-tâches particulières ont été créées pour celles-ci. A chaque type de segment sont associés des sous-espaces différents du modèle de domaine et des méthodes d'adaptation différentes. En plus des sous-espaces du modèle de domaine, il est nécessaire de restreindre les informations à celles relatives à la zone géographique pertinente. Cette dernière que nous appellerons « vue spatiale » n'est pas déterminée de la même façon pour tous les types de segments de route. Nous détaillerons cela

pour les différents types de segment dans le paragraphe sur l'adaptation. On peut voir dans la figure 22 que la tâche abstraite de recherche d'information sur un segment est spécialisée par type de navigation, puis ensuite par catégorie d'information. En analysant plus en détail les catégories d'information, on trouvera encore des spécialisations possibles pour ces tâches de recherche d'information.

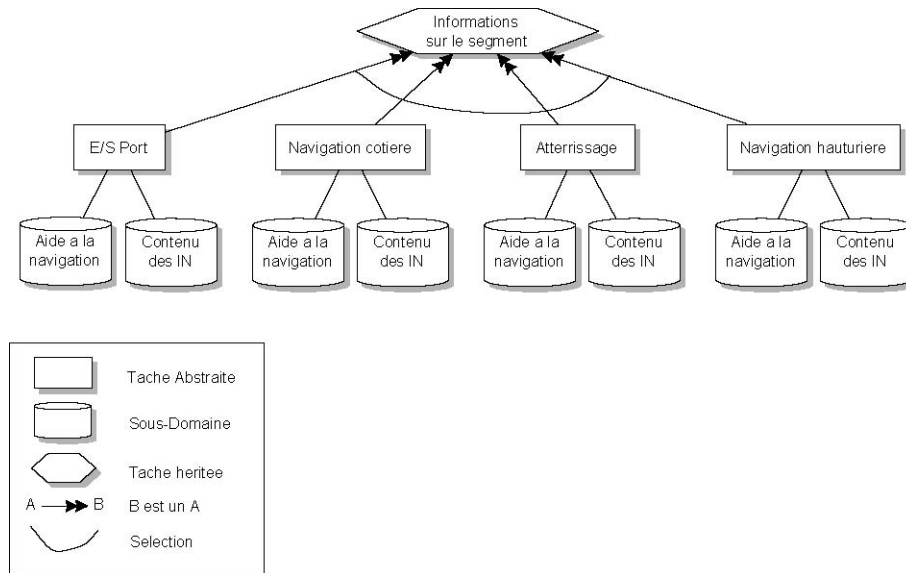


Figure 22 : Graphe d'héritage des tâches de recherche d'information.

Nous allons maintenant voir comment est décrit le modèle utilisateur pour ensuite présenter les différentes catégories d'adaptation.

3.5.2 Modèle utilisateur

La notion de stéréotype, introduite par Rich (Rich 1989), est une composante importante des modèles utilisateurs qui a été beaucoup utilisée parce qu'elle fournit un moyen simple et puissant pour l'adaptation (Kobsa 1993; Kay 1994). Les interviews ont permis de montrer que les notions de classes d'utilisateur et de tâches utilisateur - objectifs - étaient pertinentes pour réaliser des adaptations. Le modèle utilisateur se compose de stéréotypes de navigateurs et de navires, d'un modèle de tâche, d'un modèle individuel de navigateur et d'un contexte de navigation (cf. Figure 23).

- Le modèle de tâche a pour objet de mettre en œuvre les différents objectifs fournis par le système, correspondant à différents objectifs utilisateur. A chaque tâche est associé un sous-ensemble du modèle de domaine afin de restreindre l'espace d'information pour le rendre compréhensible par l'utilisateur.
- Le modèle de navire est décrit par les caractéristiques suivantes : longueur, largeur, hauteur, tonnage, vitesse moyenne du navire, catégorie de navigation (1, 2, 3, 4, 5 ou 6) qui détermine la distance maximale de navigation d'un port/mouillage - abri, type de navire (commerce, militaire, ou plaisance) qui indique des zones de navigation possible et organisation des chemises du navire s'il y a lieu. Ces propriétés regroupent

les caractéristiques physiques du navire ainsi que les informations pertinentes pour traiter la réglementation maritime.

- D'un modèle individuel de navigateur permet, pour une tâche donnée, de modifier la méthode d'adaptation par défaut (masquage au lieu d'annotation) ou les paramètres de celle-ci et de donner la profondeur minimale d'une route.

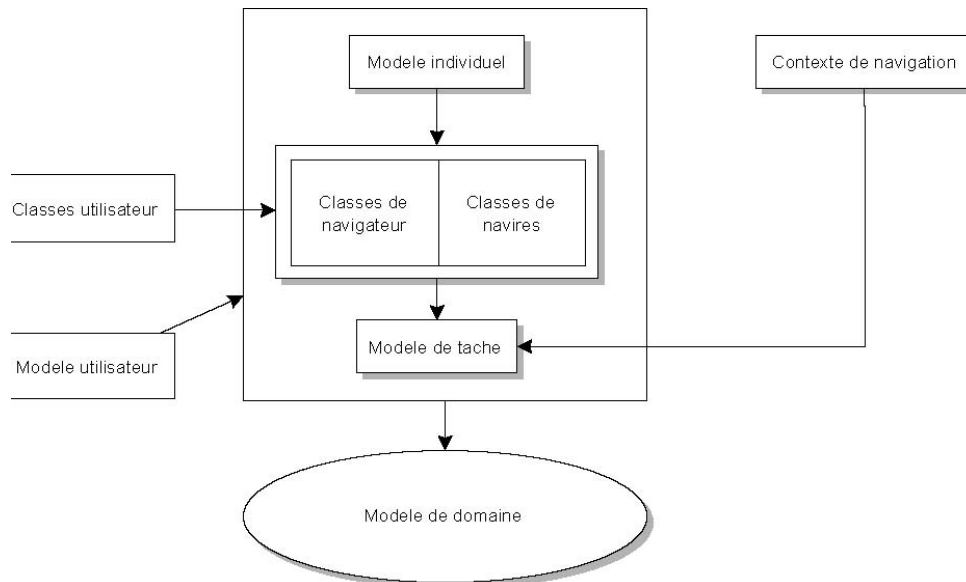


Figure 23 : Modèle utilisateur

En plus du modèle utilisateur, un contexte de navigation est utilisé pour réaliser l'adaptation. Il se compose du courant de marée, des horaires de marée, de l'heure, de la date, de la météo, de la position courante du navire (GPS ou fournie par le navigateur). Il permet donc de prendre en compte les conditions de navigation qui sont extérieures aux navigateurs et navires. Ce modèle utilisateur et ce contexte de navigation vont être utilisés pour les différentes catégories d'adaptation.

3.5.3 Les catégories d'adaptation

Nous avons utilisés deux types d'adaptation, la présentation adaptative et la navigation adaptative. La présentation adaptative est établie à partir du stéréotype de navigateur. Dans cette première version, cette adaptation ne dépend que de la catégorie du navigateur : professionnel ou plaisancier. Ce paramètre permet de choisir la présentation - conception de la page -, mais aussi le type de données puisqu'il existe des instructions nautiques pour les plaisanciers et pour les professionnels. Il est bien évident qu'il est possible d'affiner ces catégories afin par exemple de préciser les catégories pertinentes d'information pour la marine de commerce, les pêcheurs ou les militaires, etc. Pour l'instant il s'agit de montrer les effets de la présentation adaptative pour susciter des commentaires et améliorations. La présentation adaptative est indépendante des objectifs utilisateur.

La navigation adaptative est fondée à partir du modèle de tâche, du modèle de navire, du contexte de navigation et d'un modèle individuel de navigateur. Les méthodes d'adaptation utilisent un sous-ensemble de ces paramètres fonction de la tâche courante. Nous allons maintenant détailler le modèle de tâche qui est au centre de la navigation adaptative.

3.5.3.1 Adaptation pour l'objectif de recherche et création de routes

Pour pouvoir sélectionner des routes et adapter les informations présentées au couple navigateur/navire, les routes seront décrites par les propriétés suivantes :

- Un port/mouillage de départ et d'arrivée.
- Les types de navires autorisés sur cette route (plaisance, commerce, pêche, militaire, ...).
- La profondeur minimale de la route - calculé à partir des segments de route.
- La distance totale entre les deux ports/mouillages.
- Recommandée ou non.
- Les catégories de navigation autorisées sur cette route.
- La liste de couples (Points [Lat/Long], Cap compas).
- La carte sensible associée à cette route avec son tracé.
- La liste des cartes.
- Pour chaque segment de route, ligne droite entre deux « waypoints » reliés par un cap :
 - La ou les zones IN qui incluent ce segment de route.
 - La distance entre les deux « waypoints ».
 - Conditions qui rendent ce segment de route dangereux : direction et force des vents, courants, état de la mer.
 - Profondeur minimale du segment pour le zéro des cartes.
 - Le type de navigation, c'est-à-dire E/S port, navigation côtière ou hauturière ou atterrissage.

Ces propriétés permettront de retrouver les informations pour les objectifs suivants puisque les informations dont on dispose sont classées ainsi dans les IN, mais aussi d'initialiser la tâche de recherche d'information pour Info Route. L'association des segments de routes et des zones IN (Instructions nautiques - correspondantes permettra de trouver les informations utiles pour la navigation dans la bonne zone géographique pour les autres catégories d'informations.

3.5.3.2 Tâche « Affichage_Routes »

Après saisie par l'utilisateur des ports/mouillages de départ et d'arrivée, ceux-ci sont fournis en entrée de cette tâche.

- Fonction : recherche d'information et présentation des résultats
- Données en Entrée : nn port /mouillage de départ et d'arrivée
- Sous-espace d'information pertinent : routes et segments de routes.
- Données en Sortie : un ensemble de routes qui sont classées en quatre catégories : Route recommandée, Route satisfaisante, Route déconseillée, route interdite.
- Méthode d'adaptation : par défaut, annotation munie de quatre états correspondants aux quatre classes de routes.

Les critères utilisés pour l'adaptation sont les suivants pour les routes proposées :

- La catégorie de navigation : Ces routes ne seront possibles pour un navire que s'il a l'autorisation d'effectuer un transit d'une telle distance, voir les distances maximales d'un abri. Nous interpréterons ici la notion d'abri comme un port ou un mouillage cité dans un IN, permettant d'accueillir le navire considéré - tonnage, tirant d'eau, longueur, largeur, hauteur, taille du port, de la rade, de la baie, etc. les catégories de navigation du navire sont les suivantes :
 - Tout navire de longueur > 25 m, catégorie 1, 2, 3, 4, 5 qui ont pour principale caractéristique de ne pas pouvoir s'éloigner de plus d'une certaine distance d'un abri 1 : sans limite, 2 : 200 M d'un abri et distance totale du transit $= < 600$ M, 3 : 20 M, 4 : 5 M, 5 : toujours à l'abri d'une rade, étang, lac, etc.
 - Plaisance de longueur < 25 m ; 1 : sans limite, 2 : 200 M, 3 : 60 M, 4 : 20 M ; 5 : 5 M ; 6 : 2 M.
- Type de navire : les navires de commerce sauf autorisation particulière ne sont pas autorisés à naviguer dans des zones particulières, par exemple entre Ouessant et la côte. Différents types de navires : commerce, pêche, plaisance, militaire. Selon le type de navire, certaines routes sont interdites à la navigation telles que, par exemple, le FROMVEUR pour tout navire de commerce sauf autorisation spéciale, ainsi que toute la zone qui s'étend de OUESSANT à la côte.
 - Taille du navire - tonnage, tirant d'eau : elle peut obliger celui-ci à prendre des routes obligatoires, à faire des attentes sur des zones, tandis que des navires plus petits peuvent prendre d'autres routes.
 - Profondeur minimale : le navigateur peut ajouter un autre critère inclus dans ces préférences à savoir la profondeur minimale de la route > 10 m ou > 20 m - données recueillies lors des interviews.
 - Taille des navires acceptés par les ports/mouillages de départ et d'arrivée

- Les conditions de navigation pour avoir la météo, l'état de la mer, les courants, la direction et la force des vents.

Tous ces critères sont présents dans le modèle utilisateur et le contexte de navigation. Ils sont utilisés automatiquement par le système. Ce sont eux qui vont être à la source de l'adaptation qui sera ici un filtrage des routes disponibles liant deux ports/mouillages. Les règles d'appartenance d'une route aux catégories citées précédemment sont les suivantes :

- Route interdite : type de navire non autorisé, catégorie de navigation supérieure à celle de la route, profondeur d'eau minimale inférieure au tirant d'eau du navire.
- Route déconseillée : route non interdite, port/mouillage n'acceptant le navire - à cause du tirant d'eau, du tonnage ou encore danger dû au courant ou météo voir conditions liées à la route.
- Route satisfaisante : route non interdite et non déconseillée, profondeur minimale de la route supérieure ou égale à la profondeur minimale demandée par le navigateur.
- Route recommandée : route satisfaisante, fournie par les IN en tant que route recommandée, préférable à cause du courant - E/S de goulets, rades ou ports, des vents ou de l'état de la mer.

3.5.3.3 Adaptation pour la tâche de recherche d'information sur une route

Pour la recherche d'information, une tâche atomique détermine le sous-domaine pertinent qui spécifie une vue de l'hyperespace. Cette vue est encore trop importante pour être fournie à l'utilisateur. D'autres paramètres sont utilisés pour réduire cette vue. Dans le cadre des documents nautiques le critère le plus important est un critère spatial (cf. Figure 24). A partir de la position du navire, il est possible de définir une vue spatiale qui détermine le sous-espace pertinent.

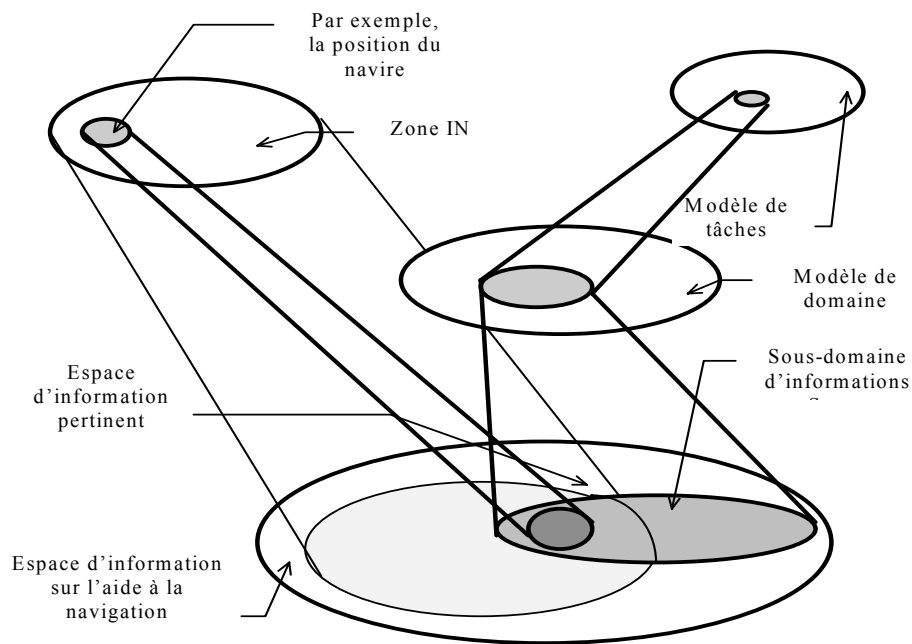


Figure 24: Vues spatiales dans SWAN.

Une vue spatiale est un espace d'information qui se compose des données appartenant aux concepts du sous-domaine pertinent dans une zone géographique donnée. Cette zone géographique peut être représentée par un polygone sur une carte nautique. Elle dépend du sous-domaine considéré et de la tâche courante. Pour certains concepts, il s'agit d'une zone des Instructions Nautiques - IN - dont la limite n'est pas définie par des coordonnées précises, mais le contenu de paragraphes des documents papiers. Pour les aides à la navigation, la zone géographique correspond à un plus petit polygone défini à priori ou résultant d'un calcul. Ce dernier est toujours inclus dans une zone IN. Une vue spatiale sélectionne les données utiles par rapport à la position du navire et à la tâche utilisateur. A partir de la position du navire, d'un segment de route ou encore de la position choisie par le navigateur, il est possible d'obtenir dans les instructions nautiques la zone IN correspondante, mais aussi celles qui l'incluent. En effet, les zones IN sont structurées par une relation de composition. Le navigateur peut ainsi accéder aux informations générales présentes dans les zones IN englobantes qui sont pertinentes pour toutes les zones incluses. Lorsqu'il s'agit des aides à la navigation - balisage, feux, signaux de brume, amers et alignement -, le calcul de la vue spatiale est fonction de la position du navire sur la route, des conditions météo et de l'heure - jour ou nuit.

La structure du document hypermédia ainsi que celle de la navigation sont fournies à l'utilisateur à l'aide du modèle de domaine et du modèle de tâche. La structure hiérarchique de notre modèle est fondée sur une relation de composition entre tâches. Cette structure forme un graphe orienté sans circuit, dont les feuilles sont les tâches atomiques et les nœuds les tâches abstraites. Immédiatement sous la racine de ce graphe se trouvent les tâches abstraites représentant les cinq objectifs offerts par le système. Une tâche possède un état qui peut prendre quatre valeurs différentes : fermer, ouverte, en attente, en utilisation. Les guides globaux et locaux sont établis à partir du modèle de tâche - tâches sélectionnées et états de celles-ci - et des catégories pertinentes d'information du modèle de domaine. Un guide global des objectifs est fourni qui permet de savoir quels sont les objectifs actifs et inactifs, et parmi

les actifs lequel est en attente. Un autre guide local est également fourni au moyen de l'état des tâches afin d'expliciter à l'utilisateur quelle est l'étape courante, mais aussi les étapes précédentes et suivantes - des tâches et/ou des catégories d'information du modèle de domaine. Lorsqu'une tâche atomique a déterminé un sous-espace d'information pertinent, le modèle de domaine correspondant est utilisé pour donner un guide local vers les différentes catégories d'information que l'utilisateur peut alors sélectionner par navigation dans ce sous-espace. Le sous-domaine concerné sert alors de repère « spatial » stable indépendamment des informations correspondantes. L'état des tâches, ainsi que la structure du sous-domaine pertinent est explicité à l'utilisateur sous forme d'hyperliens dont certains sont annotés différemment - état des tâches.

3.5.3.4 Tâches « Aide à la Navigation »

Les tâches de recherche d'information pour les informations relatives aux aides à la navigation utilisent toutes la méthode d'adaptation par annotation pour les catégories d'information. Les aides à la navigation comprennent les catégories d'informations suivantes : balisage, amer, feux, signaux de brumes et alignements (cf. Figure 25). Ils sont organisés dans un graphe qui sera visualisable par l'utilisateur.

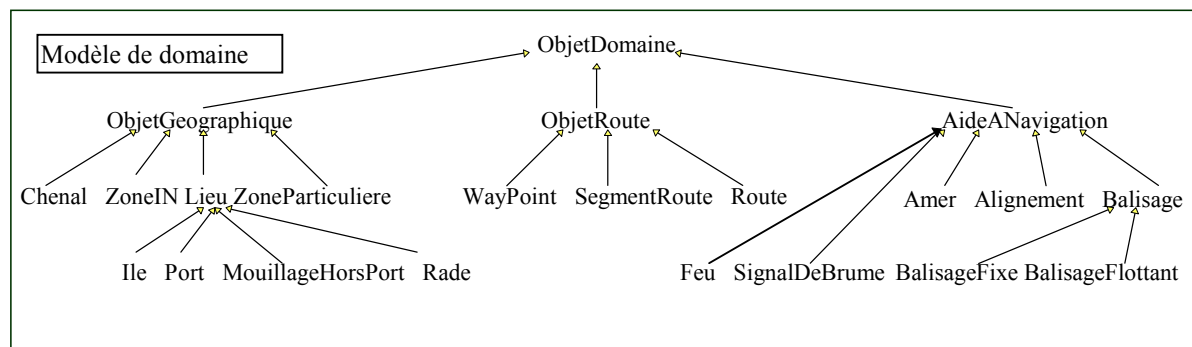


Figure 25 : Modèle de domaine de SWAN

La navigation adaptative pour ces catégories d'information est le fruit d'une opération de filtrage des informations et d'une méthode d'annotation. Pour l'annotation, deux états sont considérés : catégories pertinentes et autres. Pour chaque tâche, l'utilisateur disposera du sous graphe du domaine correspondant aux aides à la navigation dans lequel les feuilles pertinentes seront annotées d'une couleur et les non pertinentes d'une autre couleur. Cette notion de pertinence pour l'annotation prend en compte la météo - brouillard ou non -, l'heure - jour ou nuit de la manière suivante :

- Brouillard : ce sont les signaux de brume qui sont pertinents.
- Nuit : ce sont les feux et alignements avec feux qui deviennent pertinents.
- Nuit et Brouillard : ce sont les feux, les signaux de brume et alignement avec feux qui deviennent pertinents.

- Jour et sans brouillard : ce sont les balises, amers et alignements qui sont pertinents.

Le sous-espace d'information représenté par ces catégories est trop important pour être parcouru par l'utilisateur. Il contient des aides à la navigation non présentes dans la zone de navigation considérée, non visibles, ou encore non utiles parce sans rapport avec la route suivie. Il est donc nécessaire de filtrer ces informations en ne conservant que celles visibles et utiles pour la route suivie par le navigateur. Cette notion de visibilité prend en compte la portée et le secteur angulaire - zone angulaire où l'aide à la navigation est visible indépendamment de sa portée. Une vue spatiale est le résultat de la combinaison de ces trois paramètres : présence dans la zone, utile pour la route et visible. Les vues spatiales sont déterminées de manière différente pour les différents types de segment. Ces derniers vont donc induire une détermination différente de la vue spatiale. Nous devons donc considérer les trois cas suivants :

- Entrées/sortie ports : toutes les aides à la navigation visibles - portée + secteur angulaire - sont prises en compte.
- Atterrissage : une instruction nautique contient un chapitre dédié à l'atterrissage qui décrit les accès à la côte en venant du large à partir de quelques grands secteurs - un paragraphe par secteur - qui contient les amers, alignements et feux pertinents pour le secteur considéré. Il est ainsi possible de déterminer les aides à la navigation pertinentes.
- Navigation côtière : ici, la stratégie de filtrage traite différemment les diverses catégories d'information à savoir : le balisage, les feux, les amers, les alignements et les signaux de brume. Ces derniers peuvent être divisés en deux classes les aides de « proximité » et ceux qui sont éloignés que l'on range respectivement de la manière suivante : classe de « proximité » : balisage, feux sur balises, signaux de brume, classe « éloignée » : amers, amers avec feux, alignement avec ou sans feux. Pour la classe de « proximité », une zone de pertinence est définie autour du segment de route - un rectangle de la longueur du segment et d'une largeur paramétrable - par défaut 3 miles - (cf. Figure 26).

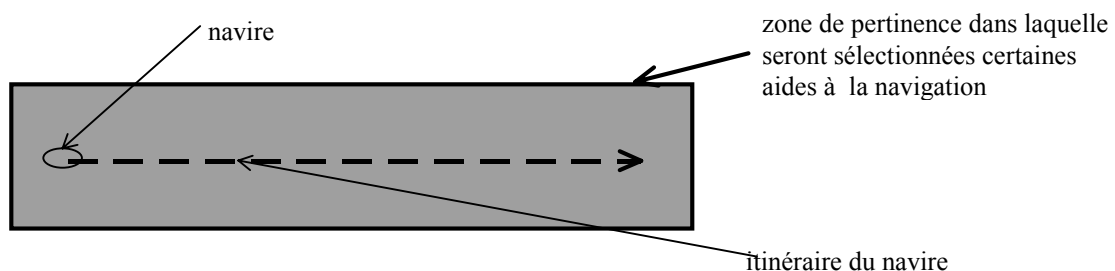


Figure 26 : Zone de pertinence.

Toutes les aides à la navigation de la classe de « proximité » présente dans cette zone sont sélectionnées et triées dans l'ordre de leur rencontre par rapport à la route pour chaque catégorie. Par contre pour les classes « éloignées », toutes les aides visibles sont sélectionnées

- portée et secteur angulaire - et triées par direction - Nord, Nord-Ouest, Nord-Est, Sud, Sud-Ouest, Ouest, Est.

3.5.3.5 Tâche « Informations IN »

On ciblera les paragraphes des zones IN correspondant à la route pour le couple navigateur/navigateur. En effet, il existe des instructions nautiques pour les professionnels et pour les plaisanciers. En fonction du navire, de son type, de sa taille, il n'est pas pertinent de fournir certaines informations. Par exemple, le navire de plaisance ne sera pas intéressé par les zones d'attentes, ni par les facilités d'un port militaire. Il en est de même pour un navire de pêche, un remorqueur. Certaines catégories d'informations sont plutôt réservées à certains navires - type et taille et ne seront donc pas présentées aux autres. Il y a donc masquage pour certaines catégories. En fonction, du type de segment de route le choix du sous-espace d'information n'est pas le même. Il est effectué de la façon suivante :

- E/S ports : partie de la zone IN relative à l'accès au port/mouillage considéré
- Atterrissage : paragraphe correspondant au secteur d'arrivée vers la zone côtière
- Navigation côtière : zone IN correspondant au segment de route.

3.5.4 Architecture du système

Le système d'information en ligne est réalisé à partir d'une architecture client/serveur, composée de trois parties ayant des rôles bien précis: un serveur HTTP, un serveur Web dynamique et un système à base de connaissance (SBC). Le serveur Web prend en charge l'interface de l'application, nous avons choisit WebObjects parce qu'il permet d'identifier et de gérer les utilisateurs par l'intermédiaire de la notion de session. Il permet aussi de générer dynamiquement des pages Web à partir de composants génériques - ou gabarits de pages. Dans la partie WebObjects de SWAN, nous disposons d'une collection de composants génériques et indépendants. En effet les liens et le parcours de lecture de ces composants sont gérés par le SBC. Notre SBC est fondé sur une logique de description, Loom qui est le « moteur » de notre application. Il contient le modèle de tâches, le modèle utilisateur, l'ontologie du domaine ainsi que les données. Le modèle de tâche étant déclaratif, il est modifiable facilement pour prendre en compte tout changement du comportement utilisateur.

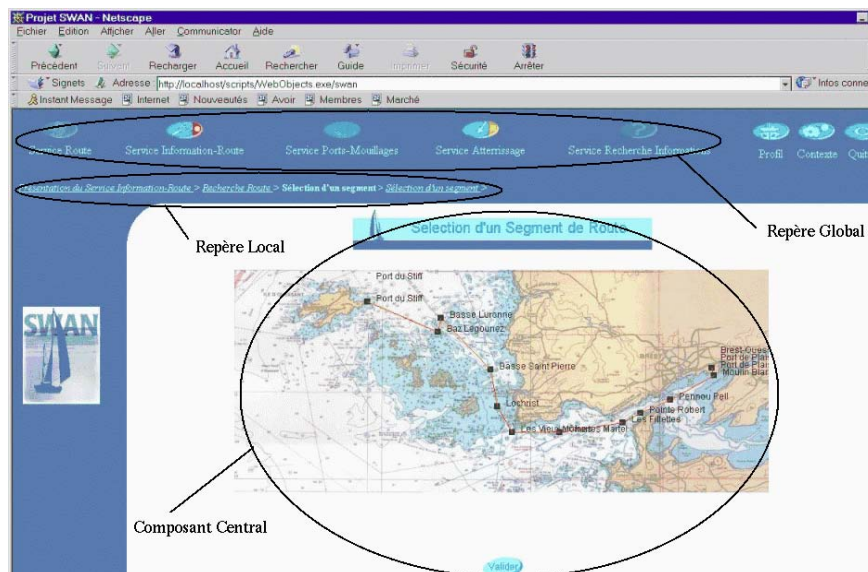


Figure 27 : Interface de SWAN

3.6 Conclusion

Ce projet propose une méthodologie de conception de systèmes d'information adaptatifs qui repose sur l'utilisation d'une structure globale - ou organisation, d'un modèle utilisateur, d'un modèle de contexte et de méthodes de filtrage. Le processus de « résolution de problème » ne nécessite pas d'expertise. Il est donc possible d'acquérir les connaissances qui vont permettre de guider l'utilisateur au travers des différentes activités qu'il doit réaliser. Le système d'information va donc prendre en charge une part plus importante par rapport à l'étape précédente : il va proposer des mécanismes de recherche d'information, une structure organisant l'accès aux informations pertinentes en fonction des tâches utilisateur et des mécanismes d'adaptation – filtrage des informations pertinentes. Les connaissances explicites de la communauté de pratiques concernée sont constituées d'un modèle de domaine qui permet d'indexer les informations, d'un modèle de tâches hiérarchiques décrivant les tâches utilisateur et organisant l'accès aux informations, d'un modèle utilisateur incluant le modèle de tâches hiérarchiques, d'un modèle de contexte et d'un modèle d'adaptation. C'est donc le modèle de tâches hiérarchiques qui permet de spécifier l'organisation. La recherche d'information se fait à l'aide de requêtes utilisant principalement un moteur d'inférence. Ces requêtes sont constituées des concepts et des relations du modèle de domaine. Par contre, le système d'information est accessible sur internet au travers d'un navigateur. Les informations ne sont réutilisables que localement. Le système d'information adaptatif est engendré dynamiquement à partir des modèles précédents.

Par rapport à l'étape précédente, les principales différences sont les suivantes :

- Le système d'information adaptatif est fondé sur un hypermédia adaptatif.
- Les besoins utilisateur sont représentés explicitement à l'aide d'un modèle utilisateur un modèle de contexte et un modèle de tâches hiérarchiques.

- Il y a une spécification explicite de l'organisation fondée sur les tâches utilisateurs qui lie la recherche d'information et l'organisation.
- Le concept de document virtuel commence à apparaître, mais n'est pas encore complètement exploité. Les éléments principaux d'un document virtuel sont déjà présents : sélection (recherche d'information), organisation et assemblage (mais de manière implicite ici).

Les mécanismes d'adaptation sont ici réparties dans différentes tâches et sous-tâches de recherche d'information. Le contrôle et la compréhension de ces mécanismes d'adaptation par un auteur et/ou concepteur n'est pas très aisée. Il en est de même pour la mise à jour de l'adaptation au fur et à mesure de l'évolution des besoins des utilisateurs, mais aussi pour l'évaluation des différentes stratégies d'adaptation. Il serait en effet intéressant de pouvoir tester différentes stratégies d'adaptation pour en évaluer les impacts. Pour cela, il serait intéressant de pouvoir maîtriser et modifier facilement ces stratégies d'adaptation

Dans l'étape suivante, le concept de document virtuel devient l'un des fondements principaux de la conception des systèmes d'information. Les documents virtuels permettent de voir la conception des systèmes d'information au travers d'un moteur de composition qui utilise trois composants principaux : la sélection ou recherche d'information, l'organisation et l'adaptation. Ces derniers amènent rapidement à monter le niveau d'abstraction pour analyser ces systèmes. En d'autres termes, on ne pense plus en termes de pages web et d'hyperliens, mais de ces trois composants comme nous le verrons au prochain chapitre. La maîtrise et la modification aisée des stratégies d'adaptation deviennent également un enjeu important.

3.7 Crédits

Les études menées dans le cadre du projet SWAN doivent beaucoup à l'équipe du SHOM participant au projet, à la société Atlantide et plus particulièrement à M. Patrick Poupon et à M. Philippe Kervella, aux différents stagiaires qui ont travaillé sur ce projet, tout particulièrement à Melle Jobard qui a notamment acquis les connaissances de domaine et enfin M. Sébastien Iksal qui a réalisé le prototype de site web adaptatif en collaboration avec des ingénieurs d'ATLANTIDE.

3.8 Bibliographie

- Boyle, C. and A. Encarnacion (1994). "Metadoc: an adaptive hypertext reading system." User Modeling and User-Adapted Interaction(4): 1-19.
- Bra, P. D. (1999). Design issues in adaptive hypermedia application development. Workshop on Adaptive Systems and User modeling on the World Wide Web, Toronto and Banff.
- Bra, P. D. and L. Calvi (1998). "Aha! an open adaptive hypermedia architecture." The new Review of Hypermedia and Multimedia 4: 115-139.
- Brusilovsky, P. (1996). "Methods and techniques of adaptive hypermedia." User Modeling and User-Adapted Interaction 6(2-3): 87-129.
- Brusilovsky, P. (2001). "Adaptive Hypermedia." User Modeling and User-Adapted Interaction 11: 87-110.

- Brusilovsky, P., L. Eklund, et al. (1998). Web-based Education for all: A tool for development adaptive courseware. Seventh International World Wide Web Conference.
- Brusilovsky, P. and L. Pesin (1994). ISIS-Tutor: an adaptive hypertext learning environment. JCKBSE'94 Japanese-CIS Symposium on Knowledge-based software engineering, Pereslavl-Zalesski, Russia.
- Brusilovsky, P., L. Pesin, et al. (1993). Towards an adaptive hypermedia component for an intelligent learning environment. Human computer Interaction. L. J. Bass, J. Gornostaev and C. Unger. Berlin, Springer Verlag. 753: 348-358.
- Conklin, J. (1987). "Hypertext: An introduction and Survey." IEEE Computer.
- Encarnação, L. M. (1995). "Adaptivity in graphical user interface: an experimental framework." Computers & graphics 19(6): 873-884.
- Fischer, G. (1990). "Communication requirements for cooperative problem solving." Informations systems 15(1): 21-36.
- Fraissé, S. (1997). Une approche systématique pour la conception de documents hypermédias. Hypertextes et Hypermédia. J. P. Balpe, A. Lelu, M. Nanrd and I. Saleh, Hermes: 249-260.
- Garlatti, S. (1997). Adaptive hypermedia for Decision Support Systems. 7th Euro Conference on Decision Support Systems and groupware multimedia electronic commerce, Brugges.
- Garlatti, S., S. Iksal, et al. (1999). Adaptive On-line Information System by means of a Task Model and Spatial Views. Second Workshop on Adaptive Systems and user Modeling on the World Wide Web, Toronto and Banff, Eindhoven University of Technology.
- Grunst, G. (1993). Adaptive hypermedia for support systems. Adaptive user interfaces: Principles and Practice. M. Schneider-Hufschmidt, T. Kùme and U. Malinowski. Amsterdam, North-Holland: 269-283.
- Hahn, W. V. and W. Wahlster (1983). Ham-ans: a german language dialog system for knowledge-based access to relational data bases. SIGART Newsletter.
- Henze, N. and W. Nejdl (2004). "A Logical Characterization of Adaptive Educational Hypermedia." New Review on Hypertext and Hypermedia (NRHM) 10(1): 77-113.
- Höök, K., J. Karlgren, et al. (1996). "A glass box approach to adaptive hypermedia." User Models and User Adapted Interaction(6).
- Iksal, S. (2002). Spécification Déclarative pour des documents virtuels personnalisables. IASC. Brest, ENST Bretagne.
- Kaheneman, D. (1973). Attention and Effort. Englewood Cliffs, Prentice Hall.
- Kaplan, C., J. Fenwick, et al. (1993). "Adaptive hypermedia Navigation Based on User Goals and Context." User Models and User-Adaptive Interaction 3(3): 193-220.
- Kay, J. (1994). Lies, Damned Lies and Stereotypes: Pragmatic approximations of users. Conference on user modeling, Hyannis, MA.
- Kay, J. and R. Kummerfeld (1994). An individualised Course for the C programming Language. Second international WWW conference "Mosaic and the Web", Chicago, IL.
- Kobsa, A. (1993). User modeling: Recent Work, Prospects and Hazards. Adaptive User Interfaces: Principles and Practice. M. Schneider-Hufschmidt, T. Kùhme and U. Malinowski. Amsterdam, North-Holland.
- Kobsa, A., D. Müller, et al. (1994). KN-AHS: an adaptive hypertext client of the user modeling system BGP-MS. 4th International Conference on User Modeling, Hyannis, MA.

- Kobsa, A. and W. Pohl (1995). "The user modeling shell system bgp-ms." User Modeling and User-adapted Interaction(4): 59-106.
- Kobsa, A. and W. Wahlster (1989). User Models in Dialog Systems, Springer Verlag.
- McTear, M. F. (1993). "User modelling for adaptive computer systems: a survey of recent developments." Artificial Intelligence Review 7: 157-184.
- Petrelli, D., A. D. Angeli, et al. (1999). A User-centered approach to user modeling. UM99, Proceedings of the Seventh International Conference. J. Kay, Springer: 255-264.
- Pohl, W. (1996). Learning about the user. User modeling and machine learning, Workshop: Machine learning meets Human-Computer Interaction, ICML'96. V. Moustakis and J. Herrmann: 29-40.
- Rich, E. (1989). Stereotypes and user modeling. user models in dialog systems. A. Kobsa and W. Wahlster. berlin, Springer verlag: 35-51.
- Rosis, F. D., B. d. Carolis, et al. (1993). User Tailored Hypermedia Explanations. INTERCHI'93 Adjunct Proceedings, Amsterdam.
- Sebillotte, S. (1995). "Methodology guide to task analysis with the goal of extracting relevant characteristics for human computer interfaces." International journal of Human computer interaction 7(4): 341-363.
- Sebillotte, S. and D. L. Scapin (1994). "From users' task knowledge to high-level interface specification." International journal of Human computer interaction 6(1): 1-15.
- Thüring, M., J. Hannemann, et al. (1995). "Hypermedia and Cognition: Designing for Comprehension." Communication of the ACM 38(8): 57-66.
- Tricot, A. and J. Nanard (1997). Un point sur la modélisation des tâches de recherche d'informations dans le domaine des hypermédias. Hypertextes et Hypermédia. J. P. Balpe, A. Lelu, M. Nanrd and I. Saleh, Hermes: 35-56.
- Tyler, S. W. and S. Treu (1989). "An interface architecture to provide Adaptive Task-Specific Context for the user." International Journal of Man, Machine and Studies(30): 303-327.
- Vassileva, J. (1996). "A Task-Centered Approach for user Modeling in a hypermedia Office Documentation system." User Models and User Adapted Interaction(6): 185-223.
- Waern, Y. (1986). Cognitive Aspects of Computer Supported tasks. New York, Wiley.

Chapitre 4. Les systèmes d'informations adaptatifs et sémantiques fondés sur les documents virtuels et le web sémantique

4.1 Préambule

De nos jours, les systèmes d'information doivent pouvoir réutiliser des ressources distribuées sur internet et s'adapter aux besoins utilisateur. Pour chaque système d'information, la masse de ressources disponibles et la variété de profils utilisateur mènent rapidement à une explosion combinatoire du nombre de versions de systèmes d'information nécessaires. Il est impossible de gérer ces systèmes d'information sans une certaine forme d'automatisation permettant de faciliter leur conception et leur maintenance. Pour toutes ces raisons, cette troisième étape marque une transition importante par rapport aux deux étapes précédentes. En effet, pour permettre la réutilisation des ressources et l'adaptation, il va être nécessaire d'utiliser des standards, de nouvelles technologies et de nouvelles méthodes de conception - document virtuel et web sémantique. La réutilisation des ressources pose les problèmes suivants : i) une augmentation de la complexité liée à la masse d'informations

disponibles et à leur hétérogénéité ; ii) la recherche d'information par mots clés sur Internet possède de fortes limitations en termes de précision et de rappel et n'est donc plus adaptée à une réutilisation par des machines de ces informations ; iii) l'interopérabilité entre les informations et les systèmes qui les réutilisent ; iv) l'automatisation ou la semi automatisation de certaines tâches - actions - du système qui doivent par exemple organiser les informations de manière cohérente - ce problème est étroitement lié à la qualité de la recherche d'information puisqu'il s'agit notamment d'obtenir les informations pertinentes au bon moment et à la bonne place dans l'organisation. La méthodologie de conception de ces systèmes d'information sémantiques et adaptatifs doit être complètement modifiée pour assurer le partage et la réutilisation des ressources, mais aussi la maîtrise et la modification aisée de l'adaptation.

Dans cette étape, nous avons proposé une méthodologie de conception fondée sur un moteur de composition flexible appelé SCARCE – SemantiC & Adaptive Retrieval and Composition Environment. Dans ce contexte, la cohérence du document produit est assurée par le savoir-faire des auteurs qui spécifient le contenu, l'adaptation et l'organisation du document à produire à l'aide de ces connaissances communes. Ce moteur de composition flexible autorise une spécification déclarative et sémantique du contenu, de l'organisation et de l'adaptation et une gestion séparée de ces derniers. Les avantages principaux de cette approche sont : i) une architecture générique qui est réutilisable dans différents contextes, ii) les connaissances explicites des communautés de pratiques sont des paramètres du moteur de composition qui sont formalisées par des ontologies. En effet, un système d'information peut être mis à jour dès que les spécifications du contenu, de l'organisation ou de l'adaptation sont modifiées. Un nouveau système d'information peut être conçu en modifiant l'une ou plusieurs des ontologies et en produisant de nouvelles spécifications. Bien évidemment, un tel procédé est limité par les principes fondamentaux sous-tendant le moteur de composition. La répartition des rôles de l'interprétation et de l'action est donc la suivante : les auteurs assurent les tâches de spécifications fonction de leur savoir-faire, les lecteurs doivent lire et comprendre le document hypermédia produit et le système d'information assure la composition adaptative des documents produits en sélectionnant les contenus pertinents et en organisant ces contenus en fonction des besoins des lecteurs. Le savoir-faire des auteurs, représenté à un niveau connaissance, est donc partageable et réutilisable par les communautés de pratiques concernées (par exemple : enseignants, journalistes, industriels, etc.). Cette répartition est bien évidemment limitée par notre capacité à acquérir ces connaissances explicites et par le rapport coût/gain de cette acquisition, des spécifications correspondantes et de la saisie des métadonnées. SCARCE est le cœur des projets ICCARS, CANDLE et KMP, mais seuls les deux premiers sont présentés dans ce chapitre :

- Le projet ICCARS (Integrated and Cooperative Computer-Assisted Reporting System) est de concevoir un système d'aide à la composition de dossiers thématiques personnalisés sur internet, c'est-à-dire de concevoir un environnement permettant aux journalistes de rédiger des dossiers thématiques adaptés à leur lectorat.
- Le Projet CANDLE (Collaborative And Network Distributed Learning Environment) est un projet européen IST (IST- 1999-11276). Il a pour objet l'utilisation d'Internet pour améliorer la qualité et réduire les coûts d'enseignement en Europe. Il s'agit de mettre en œuvre des outils de création auteur et un système de production de cours adaptés aux apprenants et à différentes approches pédagogiques. Ces cours peuvent être caractérisés comme des sortes de photocopies en ligne personnalisés.

- Le Projet KMP pour « Knowledge Management Portal » est un projet RNRT. L'objectif du projet est de construire une solution innovante de « Knowledge Management » partagé entre différents acteurs, au sein d'une même communauté de pratique. Cette solution repose sur l'analyse, la conception, la réalisation et l'évaluation d'un prototype de service web de compétences, de type document virtuel personnalisable. Ce prototype est une composante d'un portail web destiné à une communauté d'entreprises, d'institutionnels et d'organismes académiques intéressés par le domaine des télécommunications (communauté étudiée dans l'application : Telecom Valley).

4.2 Introduction

Dans cette troisième étape, nous nous sommes intéressés à des applications pour lesquels il existe au moins deux catégories d'utilisateurs : des lecteurs qui ont pour tâches principales la lecture et la compréhension d'un document hypermédia adapté à leurs besoins ; des auteurs qui possèdent des compétences et des savoir-faire leur permettant de spécifier les propriétés nécessaires à la génération de documents adaptés à ces lecteurs. Les auteurs appartiennent à une communauté de pratiques : ils partagent des connaissances communes qui leur permettent de travailler ensemble. Ce sont les connaissances explicites de ces communautés de pratiques qui sont la clé pour la conception et l'automatisation de la génération de ces systèmes d'information.

Le concept de document virtuel, qui permet la réutilisation des informations et la génération dynamique des systèmes d'informations, devient donc fondamental. Les documents virtuels sont des documents qui n'ont pas d'état persistant et pour lesquels chaque instance – document réel - est générée dynamiquement en fonction de la demande utilisateur. En d'autres termes, un document virtuel adaptatif possède un moteur de composition qui sélectionne, organise, adapte et assemble dynamiquement des ressources multimédia pertinentes. Le document réel est généré dynamiquement à partir de modèles sémantiques – ontologies – et d'une spécification. Une spécification est un paramètre du moteur de composition qui peut définir au plus les critères de sélection, d'organisation, d'adaptation et d'assemblage - présentation. Le web sémantique et ses standards permettent d'associer du sens à l'adaptation, à l'organisation et aux ressources pour améliorer la recherche d'information et permettre la réutilisation de ressources et une automatisation de la création des systèmes d'information. En effet, les moteurs de recherche sémantique constitués de métadonnées sémantiques, des ontologies correspondantes et des moteurs d'inférence, surmontent les limitations de la recherche par mots clé en associant du sens aux ressources. Les métadonnées et les ontologies correspondantes se substituent aux ressources et définissent un contexte de réutilisation de celles-ci.

Les enjeux, mentionnés ci-dessus, peuvent être notamment précisés de la manière suivante :

- La description du sens, la distribution de l'interprétation et de l'action entre l'homme et la machine et l'interopérabilité des ressources au niveau sémantique :

Ces deux enjeux ont plus d'importance qu'aux deux premières étapes parce que les ressources, leurs descriptions par des métadonnées sémantiques et les ontologies

correspondantes ne sont plus maîtrisées localement, mais peuvent être produites par une communauté géographiquement répartie. La description du sens est réalisée à partir d'un schéma de métadonnées et des ontologies correspondantes qui assurent l'interopérabilité au niveau sémantique. Il est donc d'autant plus nécessaire de coordonner et de partager les connaissances de cette communauté, les guides de bonnes pratiques et la méthodologie pour assurer une bonne qualité de cette description.

- La méthodologie de conception de ces systèmes d'information sémantiques et adaptatifs doit donc être complètement modifiée pour assurer le partage et la réutilisation des ressources et l'adaptation. Dès que l'on dispose du concept de moteur de composition permettant d'engendrer dynamiquement un système d'information, on est amené à se poser la question de sa réutilisation dans des contextes différents. Serait-il possible de réutiliser ce moteur de composition pour d'autres types de systèmes d'information ? En d'autres termes, comment assurer une bonne flexibilité et extensibilité du moteur de composition ?

Flexibilité : plus les mécanismes du moteur de composition sont paramétrables, plus il est facile de maintenir et de concevoir de nouveaux systèmes d'information. De nouveaux documents réels peuvent être générés dès que de nouveaux paramètres de la sélection, de l'adaptation, de l'organisation et de l'assemblage sont définis et/ou modifiés.

Extensibilité : à tout document virtuel sont associées des propriétés fondamentales qui caractérisent le type des systèmes d'information produits. Un moteur de composition met en œuvre ces caractéristiques. Un moteur de composition n'est réutilisable que si ces propriétés fondamentales sont adéquates aux domaines d'application envisagés. En d'autres mots, si ces principes sont compatibles avec la conception d'un type de système d'information particulier, on peut le réutiliser. Il est bien évident que ces principes fondamentaux sont aussi une limitation à la réutilisation. La conception du moteur de composition doit donc faciliter l'extension de ces propriétés.

- Organisation et acquisition des connaissances : comme le SI est fondé sur des connaissances, il est nécessaire d'avoir une méthodologie d'acquisition et de structuration des connaissances qui permettent une acquisition et une mise à jour aisée de celles-ci.

Ce chapitre est organisé autour de 6 paragraphes. Le premier propose un état de l'art sur les documents virtuels. Le second et le troisième paragraphe introduisent le web sémantique et les métadonnées relativement à notre problématique. Le quatrième paragraphe présente l'environnement SCARCE dédié à la conception de systèmes d'information adaptatifs et sémantiques au travers des projets ICCARS et CANDLER. Les deux derniers sont constitués de la conclusion et de la bibliographie.

4.3 Les documents virtuels

Comme nous le verrons plus en détail dans ce paragraphe, les documents virtuels proposent des méthodes et des techniques pour la réutilisation des informations et la génération dynamique des systèmes d'informations. Cette dernière va aussi permettre de faciliter la création et la maintenance de systèmes d'information.

4.3.1 Historique

Le document virtuel prend ses racines au sein des travaux sur les hypermédias flexibles qui ont débuté dans la fin des années 80 (Iksal 2002). Un système hypermédia flexible est un système qui propose différentes vues d'un même hypermédia en fonction de l'utilisateur ou qui génère dynamiquement l'hypermédia et le contenu de cet hypermédia au moment de la consultation. La notion d'hypermédia flexible a pour but de répondre aux besoins tels que la présentation d'une information la plus récente et/ou la présentation d'une information personnalisée. Elle se décline autour de trois approches (Oberlander and Milasavljevic 1997):

- L'utilisation des techniques venant de la recherche d'information et du filtrage afin d'évaluer la pertinence des nœuds d'un hypermédia et de modifier la présentation de l'information en fonction de cette évaluation.
- Les hypermédias adaptatifs (Brusilovsky 1996) utilisent un modèle de l'utilisateur afin de modifier l'hypermédia existant et de présenter le contenu en fonction des besoins utilisateur.
- Les hypermédias dynamiques ne modifient pas un hypermédia existant, mais le génèrent dynamiquement en fonction du type d'utilisateur ou de l'historique du discours - prise en compte du contexte.

Les hypermédias flexibles n'ont pas pour objectif la réutilisation d'informations préexistantes localement ou sur internet. Face à la masse d'informations disponibles sur Internet, il semblait intéressant de faire évoluer les hypermédias flexibles de façon à réutiliser au maximum cette connaissance. Le Document Virtuel enrichit la définition de l'hypermédia flexible en insistant d'une part sur la réutilisation d'informations disponibles sur diverses sources (le Web par exemple) et d'autre part sur les mécanismes de génération du document pour un utilisateur. Afin de définir au mieux le document virtuel en fonction des différentes définitions proposées dans la littérature, nous allons revenir sur certaines d'entre elles.

4.3.2 Définitions

Il est intéressant de noter que dès 1989 (Nanard and J. Nanard 1989), Marc et Jocelyne Nanard introduisent le concept de document conceptuel de la façon suivante : "Un document conceptuel spécifie la structure logique, la présentation, mais aussi la sémantique du contenu d'un document et l'interaction de l'usager avec des constituants de ce document. Un document conceptuel permet de plaquer à la demande une structure particulière sur un

ensemble d'informations existantes, pour en faciliter la manipulation.". Les informations manipulées par MacWeb sont des fragments (appelés grains) typés et reliés entre eux par des relations sémantiques. Le document conceptuel est en fait un ensemble de scripts qui sont instanciés afin d'offrir une vue d'ensemble à un utilisateur. Ce document conceptuel peut être considéré comme la racine du document virtuel. Elle présente le côté dynamique de la génération, la réutilisation des informations ainsi que la définition des principes de construction.

En 1993, Thomas Gruber définit les documents virtuels comme des documents hypermédia qui sont générés à la demande en fonction de plusieurs sources d'information et en réponse à une demande de l'utilisateur (le lecteur). Ces documents, une fois générés, ressemblent à des documents Web traditionnels et ne peuvent en être différenciés (Gruber, Vemuri et al. 1996). Dans cette définition, Gruber insiste sur le fait que le document virtuel est un document hypermédia et le définit en quelque sorte comme un hypermédia flexible.

En 1995, dans le cadre de leurs travaux sur les hypertextes à basés sur de la connaissance, Marc et Jocelyne Nanard (Nanard and Nanard 1995) définissent le document virtuel comme un hypertexte dont les nœuds virtuels sont des documents générés à la demande à partir de contenus existants. Le document virtuel selon eux, présente une vue de l'information en fonction des tâches de l'usager. On pressent dans cette définition le besoin de réutilisation de l'information existante pour la construction d'un hypertexte. En 1999, lors de la première conférence sur les documents virtuels, Carolyn Watters et Michael Shepherd (Watters and Shepherd 1999) définissent brièvement le document virtuel comme étant un document qui ne possède pas d'état persistant et pour lequel toutes les instances sont générées lors de la consultation. Maria Milosavljevic lors de l'introduction de la conférence (Milosavljevic, Vitali et al. 1999), précise que le document virtuel est composé des informations et des mécanismes nécessaires à la génération du document réel qui sera consulté par le lecteur. En effet, sa définition est la suivante : "'Virtual documents' are web documents for which the content, nodes or links, or all three, are created as needed. There already exist several kinds of virtual documents on the web for which the content is determined dynamically. First, a template can be used for which node contents are substituted at runtime. Second, applications, like Maple or Mathematica, can be used to generate values for one time use. Third, CGI scripts and search engines can be used to compose virtual documents from fragments of other documents for the user on demand. Fourth, metadata can be generated for summarization for users, where the extraction and summarization is done on the fly for the user. Finally, natural language generation techniques can be employed to dynamically construct virtual documents from underlying data in data or knowledge bases."

Nous pouvons en conclure que les points importants de la définition d'un document virtuel sont les suivants :

- Les informations, encore appelée fragments, sont réutilisées à partir de différentes sources.
- Le document présenté à l'utilisateur est généré dynamiquement soit en fonction du lecteur soit en fonction d'une tâche à accomplir et produit un document hypermédia appelé document réel.

- Le document virtuel correspond à un espace d'information et des mécanismes de génération.

Les Documents Virtuels ont aussi pour objectif l'adaptation à l'utilisateur. Mais, ce sont les processus de génération dynamique qui sont étudiés le plus souvent. L'adaptation n'a guère été abordée. Pourtant l'adaptation fait partie de la définition de Maria Milosavljevic et al. : "... to customize information and display to the individual users and their tasks." (Milosavljevic, Vitali et al. 1999). Un document virtuel peut être adapté à partir d'au moins un des trois critères suivants :

- Les contraintes utilisateur : ce sont des données non persistantes définies à partir d'une requête, d'un formulaire ou d'un simple hyperlien qui lui sont propres.
- Le modèle utilisateur : il s'agit d'un ensemble de données persistantes qui caractérisent un utilisateur ou un groupe d'utilisateurs particuliers. Un tel modèle peut contenir des caractéristiques sur les connaissances, les préférences, les objectifs / tâches, les centres d'intérêts, etc. d'un utilisateur (Rich 1989; Brusilovsky, Schwarz et al. 1996).
- Le contexte (Brézillon 2002; Chevrin, J. Rouillard et al. 2006) : le périphérique d'affichage, la taille d'écran, le réseau, l'historique des interactions, la localisation géographique, les multi modalités, les multicanaux, les conditions météorologiques, etc. sont des éléments qui peuvent être pris en compte pour déterminer le contexte. On fait ici l'hypothèse que le contexte est constitué d'un ensemble de propriétés qui ne sont pas dépendantes de l'utilisateur mais de contraintes externes à celui-ci. Ce contexte sera amené à jouer un rôle de plus en plus prépondérant dans la conception de systèmes d'information.

En 1999, Sylvie Ranwez et Michel Crampes définissent les documents virtuels personnalisables – adaptatifs - comme les éléments d'information et les mécanismes permettant de réaliser une construction dynamique de documents adaptatifs. Un DVP peut être considéré comme un ensemble d'éléments (ou fragments) associé à des mécanismes de sélection, de filtrage, d'organisation et d'assemblage sous contraintes, c'est-à-dire en respectant un modèle de l'utilisateur et des principes narratifs (Ranwez and Crampes 1999). Un DVP est constitué de deux parties indissociables : la spécification et la composition. La spécification, c'est l'ensemble des informations ou des paramètres nécessaires au système pour générer un document réel. Cette spécification peut se réduire aux seules contraintes utilisateur présentées ci-dessus (par exemple : résumé vidéo d'une durée de dix minutes d'un spectacle). La composition a pour objet la construction d'un document réel à partir d'une spécification. Ces deux parties sont bien évidemment fortement dépendantes l'une de l'autre : la nature même d'une spécification dépend des principes fondamentaux de la composition et des paramètres acceptés / attendus par la composition. Spécification et composition sont en quelque sorte deux faces d'une même pièce de monnaie ; elles ne peuvent être conçues l'une sans l'autre et forment un tout indissociable.

La composition varie d'un système à l'autre, néanmoins elle comporte au moins trois processus qui peuvent être séparés ou imbriqués : la sélection et le filtrage – raffinement de la sélection - des fragments, l'organisation de ceux-ci et leur assemblage. Nous allons appeler la combinaison de ces trois processus : moteur de composition. Un moteur de composition peut

être vu comme un mécanisme capable d'engendrer une quantité « quasi-infinie » de documents réels. Il peut être réutilisé pour la conception de différents systèmes d'information à partir du moment où les principes fondamentaux de la composition et de la spécification sont adéquats au système d'information à concevoir. La sélection et le filtrage – raffinement de la sélection – consistent en un mécanisme de recherche d'information permettant la recherche et la réutilisation des fragments disponibles. L'organisation consiste à associer aux fragments sélectionnés, une ou plusieurs structures cohérentes et compréhensibles par le lecteur du document réel afin d'organiser l'accès à ces fragments et donc de leur donner du sens à l'aide de ces structures. Sélection et organisation sont très dépendantes l'une de l'autre. La spécification de ces deux processus se fait donc bien souvent en même temps. En d'autres termes, leurs spécifications sont généralement couplées. L'assemblage a pour objet de présenter à l'utilisateur une interface homme/machine qui lui permette l'interaction avec le document réel et qui facilite la compréhension de celui-ci – comme par exemple la création de guides locaux et globaux de navigation.

Nous allons présenter l'étape de composition ainsi que les implications correspondantes en termes de spécification de document. La sélection, l'organisation et l'assemblage peuvent être analysés tant du point de vue de la composition que du point de vue de la spécification. Ces trois processus et leurs spécifications permettent de concevoir les systèmes d'information avec un niveau d'abstraction bien plus élevé que celui bien souvent utilisé pour des hypermédia, à savoir des nœuds, des liens et des pages web (par exemple).

4.3.3 Le processus de sélection

Le processus de sélection a pour objet de rechercher les fragments disponibles et de ne réutiliser que les fragments pertinents. Pour effectuer cette sélection, il est nécessaire de définir la nature des fragments, leur structure et leur indexation. Un point important est le choix de la granularité des fragments (nature et structure) car ce choix conditionne la composition du document, mais aussi l'indexation et donc la recherche d'information. On peut distinguer deux niveaux de granularité l'un pour la recherche et l'autre pour la réutilisation (Vercoustre, Dell'Oro et al. 1997). En effet, un fragment peut être sélectionné, mais seule une partie de celui-ci peut être réutilisée. La granularité permet de définir certaines fonctionnalités proposées, ainsi qu'une adaptation plus ou moins fine. On peut par exemple avoir un fragment par niveau de connaissance des apprenants (1^{re}, 2^e, 3^e année), ou un fragment unique constitué de trois parties, une pour chaque année. Il est bien évident que dans ces conditions le moteur de composition ne traite pas les fragments de la même façon.

Un document est composé d'un ensemble de fragments d'information qui peuvent provenir de différentes sources (Web). Le processus de sélection correspond à l'application de méthodes de recherche d'information sur un espace d'information particulier : base de données, ensemble de pages (sites) Web, etc. La sélection retient les fragments pertinents qui constitueront le contenu du document réel. C'est une étape de recherche d'information qui doit apparier une spécification de contenu avec les index des fragments. La spécification du contenu peut être obtenue de deux façons : soit un auteur spécifie l'organisation et la sélection, soit le système propose un ensemble de « règles » qui permettent au lecteur de spécifier le contenu. L'indexation des fragments peut être réalisée par de simples mots clés, un thésaurus, des vocabulaires standardisés, un modèle de domaine formalisé par une ontologie ou encore un schéma de métadonnées formalisé ou non par des ontologies. Le mécanisme de

recherche de fragment pourrait aussi utiliser des moteurs de recherche type Google avec lesquels on ne « maîtrise » pas l'indexation et la nature des fragments. Dans tous les cas, la caractérisation de ces fragments et de leur indexation ne peut pas être séparée de celle des mécanismes d'un DVP.

Dans la plupart des DVP, les concepteurs proposent une spécification de la sélection qui dépend de l'indexation des fragments, de leur stockage et des langages et/ou mécanismes de requêtes. Cette spécification est généralement couplée avec celle de l'organisation. Anne-Marie Vercoustre propose des modèles de pages dont une partie est statique et une autre dynamique. Pour la partie dynamique, l'accès aux informations est obtenu par des requêtes SQL encapsulées dans des scripts CGI (Vercoustre, Dell'Oro et al. 1997). Gilles Falquet utilise aussi une spécification de la sélection dans un langage proche de SQL permettant de retrouver l'information au sein d'une base de données (Falquet, Nerima et al. 1999). La spécification est alors mémorisée directement dans chacun des nœuds constituant le document hypermédia. Une autre solution à base de scripts a été adoptée dans le projet MacWeb (Nanard and J. Nanard 1989), le système possède un langage de script permettant de spécifier la totalité des opérations à effectuer sur les fragments pour générer le document. Il s'agit des opérations de sélection de l'information, d'organisation, de présentation et aussi d'interaction avec l'utilisateur. Les informations sont des fragments (appelés grains) typés et reliés entre eux par des relations sémantiques dans une ontologie de domaine. Ils sont donc instances d'un concept - ou d'une classe. L'auteur du document élabore des scripts afin de spécifier son document. Dans le projet Sibyl (Ranwez 2000), l'auteur peut aussi définir l'ensemble des objectifs pédagogiques à atteindre et le système sélectionne ensuite au travers de l'ontologie de domaine les concepts pertinents puis les fragments d'information correspondant à ces concepts.

Dans Myriad (Paris, Wu et al. 2004), l'utilisateur peut faire une « requête » en langage naturel – Tell me about Dennis Sidharta - qui est interprétée et formalisée pour ensuite être utilisée par le planificateur de document virtuel. Ce dernier choisit les opérateurs de plan pertinents en fonction du contexte et ces derniers effectuent la recherche d'information. Le plan spécifie l'organisation et les opérateurs la sélection. Un opérateur utilise une requête formelle et un modèle de domaine pour constituer une requête à des agents de recherche d'information. Ces agents peuvent interagir avec des bases de données externes ou locales, des moteurs de recherche de type Google et des comptes email.

Dans le cadre d'Opales (Isaac, Courounet et al. 2004), la notion de point de vue a été introduite pour représenter des communautés d'intérêts ayant un même objectif ou tâche. Chaque communauté d'intérêts partage une même annotation ou indexation des fragments multimédia. Pour chaque communauté, il est possible d'annoter à partir de formulaires dont la structure et les champs sont prédéfinis et de graphes conceptuels définis à partir d'une ontologie du domaine correspondant au point de vue choisi. Ces derniers permettent donc de rechercher des fragments et d'aider à l'annotation en utilisant les mécanismes d'inférences des graphes conceptuels. Dans les travaux de J. Geurts et al (Geurts, Bocconi et al. 2003), les fragments multimédia sont annotés par des graphes sémantiques – en RDF - à partir d'une ontologie du domaine. A partir d'une demande utilisateur, par exemple « life and work of Rembrandt », du choix du genre de présentation et de l'annotation des ressources multimédia faite sous la forme d'un graphe RDF, le système sélectionne les ressources pertinentes relativement à cette requête.

Dans certains cas, comme le projet Karina (Crampes, Ranwez et al. 2000; Ranwez 2000), le lecteur choisit parmi des objectifs prédéfinis, ceux qu'il veut réaliser. Puis le système recherche les fragments d'information permettant d'atteindre ces objectifs. Un objectif est présenté par un vecteur d'état conceptuel, c'est-à-dire un ensemble de valeurs d'indexation pondérées selon leur pertinence. Le choix du fragment se fait par rapprochement entre le vecteur d'état conceptuel de l'objectif et celui qui indexe le fragment. Une autre possibilité consiste à permettre au lecteur de définir son espace d'information et au système d'organiser cet espace par la suite (principe des moteurs de recherche).

La sélection de l'information peut aussi se limiter au filtrage de l'espace d'information de départ. Les filtres peuvent, par exemple être de type temporel, de type sémantique ou encore fonction des caractéristiques de l'utilisateur - préférences du lecteur ou niveau de connaissance. Dans le cadre du projet Karina, Michel Crampes et Sylvie Ranwez utilisent une méthode de filtrage basée sur les durées de lecture des fragments (Crampes, Ranwez et al. 2000; Ranwez 2000), car il s'agit de composer un cours en fonction du temps que peut y consacrer l'utilisateur. Dans les projets MacWeb (Nanard and J. Nanard 1989) et SWAN (Garlatti and Iksal 2000), le filtrage est qualifié de sémantique car il est fondé sur un modèle de domaine. En effet, dans SWAN, l'utilisateur commence par filtrer l'espace d'information par l'intermédiaire d'un modèle de tâches en fonction de critères géographiques. Ensuite, il parcourt les sous-ensembles pertinents du modèle de domaine pour sélectionner un des concepts. Une fois ce concept sélectionné, le système lui présente un espace d'information restreint qui est associé à ce concept en fonction du modèle utilisateur et du contexte de navigation.

Parfois, l'étape de sélection de l'information est remplacée par une étape de génération dynamique de l'information. Jean-Pierre Balpe propose un générateur de texte (Balpe 1997). Ce générateur utilise des grammaires, des dictionnaires et des représentations du monde qu'il utilise pour générer du texte. Par exemple, la phrase « [PERSONNE] [VETEMENT [COULEUR-DE-VETEMENT]] ... » combinée à des représentations telles que « des dames existent dans le monde », « les dames peuvent être vêtues de vêtements », « les vêtements peuvent avoir une couleur » et « le blanc est une couleur possible des vêtements » permet d'obtenir « une dame vêtue de blanc ... ». Le document généré n'est pas un hypermédia, mais un texte linéaire qui est différent pour chaque lecteur. Il s'agit d'une génération dynamique de document qui ne nécessite ni caractéristique du lecteur ni contraintes de la part de l'auteur.

4.3.4 Le processus d'organisation

Le processus d'organisation, appelé aussi ordonnancement (Tazi and Altawki 1999) ou planification (Ranwez 2000), consiste à associer à l'espace d'information obtenu lors de l'étape de sélection, une ou plusieurs structures cohérentes et compréhensibles par le lecteur. L'organisation peut être soit implicite c'est-à-dire le résultat d'un calcul, soit explicite comme par exemple une structure narrative, un plan, une approche pédagogique, etc.

Une structure implicite peut être un simple tri des différents fragments, comme dans le cas des moteurs de recherche, ou résulter d'un calcul plus complexe à base de scripts ou de règles. Dans le cas des tris, la structure est fonction des critères utilisés comme la pertinence estimée, la date de production du fragment, etc. Il peut s'agir aussi de règles de précedence par l'utilisation de pré-requis, ou encore de règles de proximité sémantique par la définition et

l'utilisation, par exemple, d'un modèle de domaine (Ranwez 2000). La notion de proximité sémantique revient à calculer l'éloignement entre deux concepts, par exemple "chien" et "chat" sont plus proches que "chien" et "parapluie". Dans le projet MacWeb (Nanard and J. Nanard 1989), l'auteur de l'hypermédia dispose d'un langage de scripts. Il spécifie dans le script l'organisation du document en fonction de l'objectif du document (Nanard and Nanard 1995). Ce type de structure peut être défini de deux façons, soit un auteur possède un environnement lui permettant de définir les règles de calculs, soit les règles sont établies par le concepteur de l'application ce qui réduit les possibilités d'évolutions et de mise à jour des règles.

Une structure explicite peut prendre différentes formes comme un modèle de tâche ou encore une structure narrative. La structure narrative proposée par Jean-Pierre Balpe consiste en une description du document à partir d'une grammaire, par exemple la structure d'un roman policier peut être représentée par « [ENIGME][ENQUETE][SOLUTION] » (Balpe 1997). L'organisation des fragments de texte dans le document est fonction du développement des différents éléments de la description. Saïd Tazi propose, dans le projet FORSIC, d'utiliser des règles pédagogiques définissant les différentes étapes permettant d'atteindre un objectif (Tazi and Altawki 1999). Par exemple, une règle pédagogique peut être : « Donner Exercice, Faire exercice et Vérifier exercice et Corriger exercice ». La combinaison des différentes règles pédagogiques correspond à un modèle de tâche qui décrit les étapes de résolution d'un problème ou de présentation d'une information. Dans le projet SWAN (Garlatti, Iksal et al. 1999), un modèle de tâche est utilisé pour spécifier l'organisation et la sélection. Il correspond à l'ensemble des étapes permettant à l'utilisateur d'atteindre un des objectifs disponibles comme la consultation des routes maritimes existantes ou de l'extrait des instructions nautiques correspondant à une position géographique donnée. Anne-Marie Vercoustre propose une approche fondée sur une structure logique de documents, en SGML (Vercoustre, Dell'Oro et al. 1997). Cette structure possède certains éléments statiques et d'autres qui correspondent à des spécifications utilisées pour rechercher l'information de façon dynamique (Requêtes dans des bases de données). Dans certains cas, l'auteur ne dispose pas d'un langage de description de document. Gilles Falquet et al proposent l'utilisation d'un schéma de nœuds (Falquet, Nerima et al. 1999). Chaque nœud possède une expression d'ordonnancement. L'objectif des schémas de nœuds est de représenter une base de données. Chaque schéma correspond à une collection de la base. Le document produit correspond à une vue de la base.

Dans Myriad (Paris, Wu et al. 2004), le planificateur de document virtuel choisi le plan pertinent en fonction du contexte. Ce dernier peut être constitué au plus – dépend de l'application - d'un modèle utilisateur, d'un modèle de domaine, d'un modèle de tâche, d'un modèle de périphérique d'affichage et d'un modèle de discours qui contient un historique des requêtes de l'utilisateur. Le plan spécifie l'organisation et les opérateurs qu'il contient la sélection.

Dans les travaux de J. Geurts et al (Geurts, Bocconi et al. 2003), plusieurs ontologies sont utilisées pour spécifier une organisation définie à partir d'une ontologie des types de discours – par exemple : « biographie et CV » constitué de plusieurs unités narratives – et d'une ontologie des agents - personnes. Chaque unité narrative détermine un agent et des contenus relatifs à ce dernier. A l'exécution, une organisation ou parcours structuré, appelée « structured progression » et muni des contenus pertinents est produite. Comme on verra dans le paragraphe suivant au sujet de l'assemblage, le calcul de l'organisation est réalisé en deux

étapes la première décrite ici et la seconde qui transforme ce parcours structuré en fonction du type de média choisi par l'utilisateur.

Les règles d'organisation définissent la structure globale du document, elles sont établies de manière à faciliter la compréhension du document par le lecteur. Elles sont par conséquent dépendantes du niveau de granularité adopté pour les fragments d'information. L'organisation sert de support pour la construction de la navigation et la compréhension. En effet, cette dernière est fortement liée à l'organisation et à l'orientation dans un document.

4.3.5 Le processus d'assemblage

Ce processus concerne toutes les opérations nécessaires afin de rendre le document plus facile à lire et à comprendre. Le système peut par exemple, ajouter des transitions entre les fragments, appelées aussi liants (Ranwez 2000) pour assurer la cohérence d'un ensemble de textes associés, ou encore effectuer des opérations sur les fragments, fondées sur de la vidéo (Ranwez 2000). Pour Rekik, l'auteur spécifie les opérations que le système peut effectuer sur les fragments (rotation pour une image, différents systèmes d'unités pour un graphique, etc.) (Rekik, Vanoirbeek et al. 1999). Le processus d'assemblage concerne aussi la représentation de la structure globale du document ainsi que la représentation visuelle du document. La représentation de la structure globale a pour but de faciliter la compréhension du document par le lecteur, mais aussi de l'aider à se repérer au sein du document. L'auteur peut spécifier cette représentation en fonction des possibilités offertes par le système : graphe, repères de navigation globaux et locaux, etc. Dans SWAN, nous proposons un repère de navigation global qui permet de situer l'utilisateur au sein de l'ensemble des services, ainsi qu'un repère de navigation local qui décrit les étapes déjà réalisées, l'étape en cours ainsi que les étapes à venir dans le service courant (Garlatti, Iksal et al. 1999). Il peut intervenir aussi sur les règles de présentation portant sur la typographie, ou l'ajout d'annotations utiles pour orienter le lecteur. Le langage de scripts proposé dans MacWeb (Nanard and J. Nanard 1989) est utilisé aussi afin de créer des règles de présentation. Ensuite l'auteur peut associer ces règles comme propriétés des différents objets.

Dans le système Tiddler (Paris, Wu et al. 2004), deux bibliothèques de plans sont utilisées : l'une pour l'organisation et la sélection : les opérateurs de discours et l'autre pour l'assemblage : les opérateurs de présentation. Il est bien évident que le choix des opérateurs de présentation dépend de celui des opérateurs de discours. Ce dernier est fonction du contexte qui peut contenir un modèle du périphérique d'affichage. Dans les travaux de J. Geurts et al (Geurts, Bocconi et al. 2003), l'utilisateur choisit un type de média de sortie : document imprimé, présentation hypermédia interactive, multimédia non interactif, diaporama, etc. L'organisation et le contenu, appelés parcours structuré, sont transformés en fonction du type de média pour produire un document structuré. Ce dernier est défini à partir de deux ontologies : l'ontologie des documents structurés et l'ontologie des types de média. On peut ici remarquer que le calcul de l'organisation du document réel est faite en deux étapes : l'une lors du calcul du parcours structuré et la seconde pour le calcul du document structuré. Ce dernier joue deux rôles : la transformation du parcours structuré et la présentation du document réel.

4.3.6 Conclusion

Les documents virtuels proposent des mécanismes de génération dynamique de documents (notamment hypermédia) associés à des espaces d'information à partir d'une demande utilisateur. Les informations sont recherchées puis réutilisées pour engendrer le document réel produit pour l'utilisateur courant. Peu de travaux ont porté sur l'adaptation et la gestion du contexte qui deviennent de plus en plus importants pour la conception des systèmes d'information du futur. Comme nous avons pu le voir, la génération est fondée sur trois mécanismes fondamentaux : la sélection qui utilise un mécanisme de recherche d'information permettant la réutilisation des informations, l'organisation et l'assemblage. Ces trois mécanismes permettent de mettre en évidence trois aspects fondamentaux des documents produits : la recherche du contenu – sélection, la structure globale du document produit – l'organisation et la présentation – assemblage. Ces trois aspects ne sont pas indissociables, mais il est ici possible de les gérer séparément même s'il est nécessaire de les concevoir ensemble. Cette séparation permet aussi aux concepteurs de les traiter à un niveau d'abstraction plus élevé. Au lieu de penser en termes de page web – contenu - de liens de navigation et ou de présentation des pages, il est plus aisé de réfléchir sur les principes et/ou le sens du contenu, de l'organisation et de la présentation. Il est ainsi plus aisé de traiter des besoins utilisateur à des niveaux d'abstraction plus facilement interprétables par les utilisateurs, les concepteurs et les auteurs. On le voit bien dans les travaux de J. Geurts et al (Geurts, Bocconi et al. 2003), de M. Crampes et S. Ranwez (Ranwez and Crampes 1999; Crampes, Ranwez et al. 2000; Ranwez 2000) – Sybil et Karina par exemple, dans Myriad (Paris, Wu et al. 2004), dans Opales et Saphir (Nanard and J. Nanard 2004), Swan (Garlatti and Iksal 1999). Ces différents concepteurs utilisent des principes narratifs, des structures rhétoriques, générateurs de plans, des spécifications par intentions, etc. qui sont bien plus proches du niveau de compréhension des auteurs/concepteurs que les notions de liens et de nœuds⁸. En d'autres termes, on peut exprimer ces mécanismes à un niveau sémantique et les représenter également à un niveau sémantique, notamment par des ontologies.

Pour assurer la recherche d'information sur Internet, la représentation de ces mécanismes à un niveau sémantique, il va être nécessaire de faire appel au web sémantique et aux métadonnées pour concevoir ces documents virtuels.

4.4 Le web sémantique

Dans son article de 2001, Tim Berners-Lee (Berners-Lee, Hendler et al. 2001) propose une extension du web appelé le web sémantique. Il l'introduit de la manière suivante : “ a new form of web content that is meaningful to computers will unleash a revolution of new possibilities”. Dans cette nouvelle vision du Web, encore appelé web de troisième génération, ce dernier est vu comme un espace virtuel d'échange de ressources entre êtres humains et machines permettant une exploitation de grands volumes d'informations et de services variés. Ce web sémantique doit devenir un moyen très puissant de coopération entre les êtres humains et les ordinateurs présents sur le réseau doivent pouvoir coopérer. Ils seront capables

⁸ Le propos est ici un peu exagéré puisque la conception des hypermédia a souvent été fondé sur des modèles qui permettent d'avoir un niveau d'abstraction bien plus élevé que celui des nœuds et des liens.

d'analyser toutes les ressources disponibles sur internet et de rechercher les informations pertinentes pour les utilisateurs. Ces informations seront ensuite traitées et présentées aux utilisateurs.

Le web actuel peut être considéré comme syntaxique, dans le sens où la structure des ressources est bien définie, mais que son contenu reste pratiquement inaccessible aux traitements par des machines. En d'autres termes, les contenus ne possèdent aucune sémantique interprétable par ces dernières ; ils ne sont interprétables que par des humains. Cette nouvelle génération de web ou web sémantique a pour objet de dépasser cette limitation. L'accès automatique aux ressources est fondé sur une sémantique interprétable par les machines et des heuristiques et/ou des raisonnements utilisant celle-ci. Une sémantique explicite des ressources associée à des théories sur des domaines (ontologies) permet de proposer de nouveaux services qui automatisent ou semi automatisent certaines tâches - actions. Les ressources seront plus aisément retrouvées, réutilisées, échangées et recomposées par l'homme et/ou par la machine, grâce à cette représentation sémantique des contenus. Les utilisateurs seront déchargés d'une partie de leurs tâches (recherche d'information, recomposition des ressources retrouvées) grâce aux capacités accrues des machines à accéder et à raisonner sur les contenus des ressources. Ce web sémantique peut être vu comme la mise en œuvre d'une nouvelle répartition de l'interprétation et de l'action entre l'homme et la machine qui est directement liée à celle de l'attribution du sens (au contenu) et leur traitement contrairement au Web que nous connaissons aujourd'hui.

Le web sémantique est constitué d'une architecture ou infrastructure munie de standards (langages, formats de ressources, etc.) afin d'assurer l'interopérabilité à différents niveaux (adressage, syntaxique et sémantique) et d'applications sans lesquelles cette vision ne peut exister. Au niveau sémantique, des connaissances formalisées, en plus du contenu informel actuel du web, doivent être utilisées. Il n'existe pas actuellement de consensus sur le ou les niveaux de formalisation nécessaire. Mais, la recherche des ressources, leur réutilisation, et leur recomposition à l'échelle du web ou d'un sous ensemble de celui-ci engendrent de nouveaux problèmes et de nouveaux défis : changement d'échelle dû à la quantité de ressources disponibles, la nécessité d'assurer l'interopérabilité, la standardisation, la diversité des usages, la distribution et l'impossibilité d'assurer une cohérence globale. Relever tous ces défis ne se fera qu'en ayant des niveaux de formalisation ayant un rapport coût/gain satisfaisant. En effet, plus on veut obtenir de services plus il faudra décrire finement les ressources au niveau sémantique, mais plus le coût sera élevé.

La première contribution principale du Web d'origine a été de rendre des documents disponibles au travers d'Internet, c'est-à-dire, en séparant la présentation d'un document de sa localisation. Ensuite, l'arrivée d'XML a permis de séparer la structure d'un document de sa présentation. Ce qui signifie qu'un document peut être perçu comme un conteneur de données dont la présentation peut varier selon l'utilisation. Finalement, l'idée du Web Sémantique consiste à séparer le sens de la structure, c'est-à-dire, étendre l'idée du document avec de la sémantique manipulable par des machines (Berners-Lee, Hendler et al. 2001).

4.4.1 L'architecture du web Sémantique

Le Web Sémantique est constitué d'une infrastructure illustrée par l'architecture présentée dans la figure 28. Il s'agit d'une vision simplifiée de cette dernière qui est composée de trois

niveaux : le niveau adressage (nommage) qui permet de définir l'adresse d'une ressource et ainsi de la retrouver, le niveau syntaxique qui permet la description de la structure des ressources et le niveau sémantique permet la description sémantique des contenus de ressource à l'aide de RDF (Ressource Description Framework) et de RDF schéma ou d'ontologies, ainsi qu'avec des moteurs d'inférence qui permettront de raisonner sur ces descriptions sémantiques. Ce niveau constitue l'extension nécessaire pour réaliser le web sémantique. Nous allons maintenant présenter ces trois niveaux.

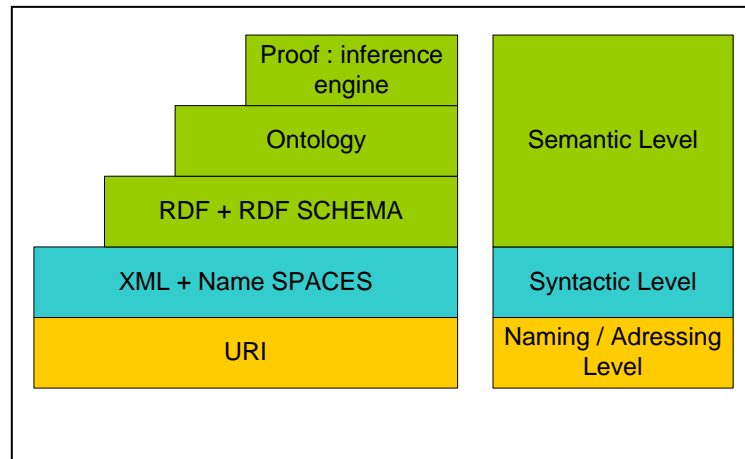


Figure 28 – Architecture du Web Sémantique

4.4.1.1 Le niveau adressage

Le World Wide Web de première génération repose sur le concept d'URI (Uniform Resource Identifier) et d'URL (Uniform Resource Locator). Toutes les ressources disponibles sur Internet sont identifiées par une URI (<http://www.w3.org/Addressing/>). Une URL identifie de manière unique et non ambiguë chaque ressource du Web, comme une page, une image, etc. Les URLs sont utilisés pour référencer des ressources Web, par exemple à l'aide du protocole HTTP. A partir du moment où toutes les ressources sont identifiées par une URL, tout le monde peut adresser/nommer une ressource sans avoir besoin de lui assigner un nouvel identifiant. Les langages associés aux ressources web tels que HTML et PDF sont des langages qui concernent la représentation des ressources. Cette première génération proposait une séparation de la présentation d'une ressource et de sa localisation par une URL.

4.4.1.2 Le niveau syntaxique

Vers la fin des années 1990, la deuxième génération du web a fait son apparition. Elle proposait une séparation entre la présentation et la structure. Le niveau syntaxique est le niveau de structuration des ressources. La spécification de la structure logique des ressources repose sur XML qui propose une infrastructure d'échange de ressources structurées. XML (eXtensible Markup Language, <http://www.w3.org/XML/>) est un métalangage qui permet de définir d'autres langages. Ce dernier est un sous-ensemble de SGML (ISO 8879). Les langages que définit XML sont des langages de structuration des ressources. XML permet la description de la structure des ressources par l'intermédiaire d'une DTD (Document Type Definition). Cette dernière définit les éléments constitutifs d'un type de ressource : la

structure syntaxique (un arbre) d'un type à l'aide d'une grammaire. Elle permet donc à toute application de vérifier la conformité d'une ressource à la DTD et donc de manipuler et transformer celle-ci lorsqu'elle est conforme à la DTD. XML permet donc d'assurer une interopérabilité au niveau syntaxique. Les points forts d'XML en plus de sa simplicité sont :

- Son indépendance quand à la plateforme utilisée.
- Son exploitation possible par un système informatique.
- La séparation de la présentation et du contenu.
- C'est un langage de description facilement extensible et personnalisable en fonction des besoins des applications.
- La gestion de la cohérence grâce aux DTD.

L'avantage d'XML est la possibilité de personnaliser la présentation des ressources en utilisant par exemple XSL (XML Stylesheet Language) qui permet de transformer automatiquement un fichier XML en une page HTML qui est consultable via un navigateur Internet, en une fiche de carnet d'adresses ou dans un autre format de présentation de l'information.

Par contre, XML n'a pas de sémantique formelle permettant l'interprétation par une machine. Le sémantique web vient donc compléter cette infrastructure par un niveau sémantique qui va permettre aux machines d'interpréter les contenus des ressources et proposer de nouveaux services.

4.4.1.3 Le niveau sémantique

Il est nécessaire de rendre les ressources interprétables par une machine en associant des descripteurs sémantiques à celles-ci. Un langage peut être interprété par un ordinateur s'il a une sémantique, c'est-à-dire si ces symboles et sa structure se réfèrent à un modèle sous-jacent qu'il soit implicite ou explicite. La sémantique n'existe qu'en relation avec quelque chose. En logique, la sémantique est fondée sur une théorie des modèles qui est composée d'un ensemble de structures décrivant des états possibles du monde. Si le système formel correspondant, muni d'un langage formel, d'une théorie axiomatique et d'une théorie des modèles, possède les propriétés de complétude et de correction, on peut garantir que toutes les déductions faites par le moteur d'inférence correspondant à la théorie axiomatique « préserve » la sémantique induite par la théorie des modèles. En d'autres termes, la machine sera capable « d'interpréter » les descripteurs sémantiques associés aux ressources qui utilisent le langage formel correspondant à ce système formel. La description sémantique du contenu des ressources est réalisée à l'aide de RDF et de RDF schéma ou des langages de description d'ontologies tel qu'OWL (<http://www.w3.org/2004/OWL/>, *Ontology Web Language*), ainsi qu'avec des moteurs d'inférence qui permettront de raisonner sur ces descriptions sémantiques.

Nous allons dans un premier temps présenter ces langages qui sont des standards du W3C. Comme nous le verrons plus tard les descripteurs sémantiques RDF permettent l'indexation

sémantique des ressources par des métadonnées et/ou des annotations (Charlet, Laublet et al. 2003; Baget, Canaud et al. 2005; Charlet, Laublet et al. 2005).

4.4.1.3.1 RDF (Resource Description Framework) et RDF Schéma

Ce sont des langages de description de métadonnées et/ou d'annotations au niveau sémantique qui permettent l'interopérabilité – au niveau sémantique - entre applications pour l'échange, le partage et la réutilisation de ressources non interprétable pour la machine (<http://www.w3.org/RDF/>).

RDF est un modèle relationnel simple. Une déclaration RDF est constituée d'un triplet (Objet, Attribut, Valeur), dont chaque membre peut être un littéral ou une ressource. D'un point de vue logico linguistique, ce triplet peut aussi être interprété de la manière suivante : (Sujet, Prédicat, Objet) ou encore Prédicat (Sujet, Objet), (comme par exemple <http://music.fi/pieces#finlandia>, creator, <http://composer.org/Sibelius>). Le modèle de données RDF est formellement défini par :

- Un ensemble appelé Ressources.
- Un ensemble appelé Littéraux.
- Un sous-ensemble de Ressources appelées Propriétés.
- Un ensemble appelé Déclarations, dont chaque élément est un triplet (prédicat, sujet, objet) « prédicat » est une propriété (membre de Propriétés), « sujet » est une ressource (membre de Ressources) « objet » est soit une ressource soit un littéral (membre de Littéraux).

La spécification de la syntaxe RDF est fondée sur XML, mais XML n'est pas le seul langage possible pour exprimer du RDF. En effet, le formalisme de RDF consiste en un ensemble de triplets. Ce n'est pas une sérialisation d'un arbre syntaxique XML, ou les branches devraient être présentées dans l'ordre spécifié par une DTD XML. Une description RDF est un graphe orienté étiqueté dont la syntaxe généralement est fondée sur XML.

RDF permet d'associer à une ressource un ensemble de triplets qui ne sont que des relations binaires dont la sémantique n'est pas définie. La sémantique de ces relations n'est définie que par un Schéma RDF ou une ontologie. RDF Schéma (RDFS) est une extension de RDF permettant de décrire les concepts et les relations utilisés dans les descriptions et de définir des contraintes de type sur les objets et les valeurs des triplets. RDF Schéma peut être vu comme un modèle orienté objet pour le web (<http://www.w3.org/TR/rdf-schema/>). Le rôle de RDFS est de définir un vocabulaire partagé par une communauté d'individus dont chaque concept et chaque relation possède une sémantique unique.

Mais, RDF schéma définit les propriétés en termes de classes de ressources auxquelles elles s'appliquent au lieu de définir les classes en termes de propriétés que les instances possèdent. En utilisant une approche de type RDF, chacun peut facilement ajouter des propriétés additionnelles à une ressource sans avoir besoin de redéfinir la description originale de ces classes. Un des avantages de cette approche centrée propriété est qu'il est très facile à chacun de dire ce qu'il veut au sujet d'une ressource. C'est l'un des principes de l'architecture

du Web (Berners-Lee 1998). Sinon, il serait nécessaire de définir une nouvelle propriété dans une classe et de la partager avec les autres. En d'autres termes, une ressource n'est pas une instance d'une classe d'un RDF Schéma. Une ressource est décrite par un ensemble de triplets RDF, chacun pouvant s'appliquer à des classes décrites dans RDF Schéma. Il permet de donc de combiner des méta descriptions de différentes parties du Web en un seul graphe. RDFS est limité en termes de pouvoir d'expression. Il est souvent nécessaire de pouvoir exprimer des contraintes supplémentaires : cardinalités min et max, contraintes entre propriétés, etc. Les ontologies dans le cadre du web sémantique sont des extensions de RDFS.

4.4.1.4 Les ontologies

Les ontologies sont apparues au début des années 90 en intelligence artificielle et plus particulièrement dans la communauté ingénierie des connaissances pour l'acquisition des connaissances dans les systèmes à bases de connaissances (Charlet 2002). Les travaux sur le raisonnement à partir de modèles développés dans les années 80 ont montré tout l'intérêt de distinguer les connaissances relatives au domaine selon leur nature et de raisonner sur ces modèles multiples en exploitant leurs spécificités. C'est cette mise en exergue des connaissances de domaine qui a amené la communauté en ingénierie des connaissances au concept d'ontologie – ensemble des objets reconnus comme existants dans le domaine (Charlet 2002). Il existe bien évidemment de nombreux types différents d'ontologies :

- Les ontologies pour les sciences (biologie, physique, électronique, etc.).
- Les ontologies pour le commerce (représentation des produits, etc.).
- Les ontologies culturelles.
- Les ontologies de métadonnées.
- Les ontologies pour des phénomènes dynamiques (tâches, services, processus, etc.).
- Etc.

Derrière cette notion d'ontologie, il y a aussi l'idée de réutilisation, c'est-à-dire qu'une ontologie de domaine serait indépendante des tâches réalisées par le système à base de connaissances. Par conséquent, une ontologie de domaine serait réutilisable dans d'autres contextes pour d'autres applications. Mais cette réutilisation reste très limitée car une ontologie de domaine dépend des tâches du système. L'ontologie est actuellement l'un des concepts clés du web sémantique puisqu'elle va nous permettre de représenter des connaissances partagées par une communauté pour décrire et indexer des ressources.

Le terme "ontologie" possède de nombreuses définitions (Charlet 2002; Charlet, Bachimont et al. 2005). J. Charlet propose une première définition de l'ontologie qui est la suivante :

- Ensemble des objets reconnus comme existants dans le domaine. Construire une ontologie c'est aussi décider de la manière d'être et d'exister des objets.

Thomas Gruber (Gruber 1993) définit une ontologie de la façon suivante :

- Une ontologie est une spécification formelle explicite d'une conceptualisation partagée.

Les attributs "formelle" et "explicite" signifient qu'une ontologie permet une interprétation automatisée par la machine de la conceptualisation. Il s'agit d'un ensemble de définitions, de primitives, de représentation de connaissance spécifique au contenu : classes, relations, fonctions et constantes d'objet. L'ontologie est donc le fait de conceptualiser l'univers du discours et les relations pertinentes dans cet univers. Les conceptualisations partagées incluent les espaces conceptuels de modélisation des connaissances de domaine ; les protocoles spécifiques au contenu pour la communication entre agents inter opérants ; et les accords concernant les représentations de théories de domaines particuliers. Dans le contexte de partage de connaissance, les ontologies sont spécifiées sous la forme de définitions de vocabulaire de représentation. Un exemple simple est la hiérarchie, spécification de classes et de leurs relations de subsomption.

Alexander Maedche et al. donnent une définition formelle d'une ontologie (Maedche, Staab et al. 2001) :

Définition 1 : Soit un langage logique L ayant une sémantique formelle dans laquelle des règles d'inférence peuvent être exprimées. Une ontologie abstraite est une structure $O = (C, \leq_C, R, \sigma, \leq_R, IR)$ consistant en :

- Deux ensembles disjoints C et R dont les éléments sont respectivement appelés Concepts et Relations.
- Un ordre partiel \leq_C sur C, appelé hiérarchie de concepts ou taxinomie.
- Une fonction $\sigma : C \times C$ appelée signature.
- Un ordre partiel \leq_R sur R, où $r_1 \leq_R r_2$ implique $\sigma(r_1) \leq_{C \times C} \sigma(r_2)$, pour $r_1, r_2 \in R$, appelé hiérarchie de relations.
- Et un ensemble IR de règles d'inférence exprimées dans le langage logique L.
- La fonction $Dom : R \rightarrow C$ avec $Dom(r) := \Pi_1(\sigma(r))$ donne le domaine de r, et la fonction $Range : R \rightarrow C$ avec $Range(r) := \Pi_2(\sigma(r))$ donne son échelle de valeurs.

Définition 2 : Un lexique pour une ontologie abstraite $O = (C, \leq_C, R, \sigma, \leq_R, IR)$ est une structure $Lex := (S_C, S_R, Ref_C, Ref_R)$ qui consiste en :

- Deux ensembles S_C et S_R dont les éléments sont appelés signes (entrées lexicales) respectivement pour des concepts et des relations.
- Et deux relations $Ref_C \subseteq S_C \times C$ et $Ref_R \subseteq S_R \times R$ appelées affectation de référence lexicale respectivement pour des concepts/reliations.
- A partir de Ref_C , nous définissons pour $s \in S_C$,
- $Ref_C(s) := \{ c \in C \mid (s, c) \in Ref_C \}$

- Et pour $c \in C$,
- $\text{Ref}^{-1}(c) := \{ s \in S_C \mid (s, c) \in \text{Ref}_C \}$
- Ref_R et Ref_R^{-1} sont définies de la même manière.

Définition 3 : Une ontologie (concrète) est une paire (O, Lex) où O est une ontologie abstraite et Lex un lexique pour O .

En ce qui concerne les langages de définition d'ontologies, le Web Sémantique s'oriente vers le langage OWL. OWL (Ontology Web Language, <http://www.w3.org/2004/OWL/>) est un langage de définition d'ontologie utilisé pour la publication et le partage d'ontologies sur le Web. OWL est dérivé de DAML+OIL et construit sur le langage RDF/RDFS. DAML+OIL est un langage de définition d'ontologies résultant de la fusion de deux projets :

- DAML : DARPA Agent Markup Language est un langage de définition d'ontologies élaboré en raison du manque de sémantique des documents écrits en XML (<http://www.daml.org/>). Le projet consiste à créer un langage pour le Web Sémantique ainsi que les outils nécessaires à l'utilisation de ce langage.
- OIL : Ontology Inference Layer (<http://www.ontoknowledge.org/oil/>) est une proposition Européenne de standard pour l'expression et la gestion d'ontologies. Il est fondé sur XML, RDF/RDFS et les logiques de description, et veut être extensible et simple. Il propose aussi des mécanismes de raisonnement sur les ontologies. OIL fait partie du projet On-To-Knowledge (<http://www.ontoknowledge.org/>) sur la gestion de connaissance.

4.4.1.5 Logique, Preuve et Confiance

Une ontologie définit les concepts et les relations d'un domaine particulier, mais n'explique pas comment les utiliser et/ou les manipuler. Le raisonnement sur les connaissances est un point important du Web Sémantique. Le raisonnement est réalisé grâce à des moteurs d'inférence. En raison du passage à l'échelle nécessaire sur internet, ces moteurs d'inférence doivent et devront être efficaces (Baget, Canaud et al. 2005). Il est en effet nécessaire de pouvoir traiter de très grandes quantités de concepts, de relations et de faits – des milliards de pages web sont indexées par Google. Afin de retrouver et rechercher l'information sur internet, il est également nécessaire de proposer des langages de requêtes évolués qui utiliseront les mécanismes d'inférence des environnements de gestion de connaissances. Mais la encore, la puissance d'expression de ces langages est généralement contradictoire avec performance. On le voit bien avec le projet KAON qui a largement diminué le pouvoir de son langage de requête par rapport à Ontobroker afin d'être capable de traiter efficacement de très nombreux concepts, relations et faits.

4.4.2 Les métadonnées

D'une certaine manière, on peut dire que les métadonnées prennent racine dans les travaux de Sir Anthony Panizzi. En effet, c'est en 1841 qu'il publia ses fameuses 91 règles pour les besoins du catalogue du British Museum. Cette publication marque l'acte de

naissance du catalogage moderne. Les métadonnées seront dans ce cas vues comme une extension des pratiques de catalogage bibliographique traditionnel. Jusqu'en 1995, les métadonnées étaient un terme surtout utilisé par des communautés impliquées dans la gestion et l'interopérabilité de données géo spatiales et dans la gestion de données et la conception, la maintenance de systèmes.

Dans un contexte traditionnel ou sur Internet, un point clé des métadonnées est de faciliter et d'améliorer la recherche d'information (Prié and Garlatti 2005). Les métadonnées sont un moyen pour rendre la recherche d'information plus efficace et plus ciblée. Elles permettent un contrôle plus précis des outils de recherche d'information. Les métadonnées fournissent aux moteurs de recherche des informations spécialisées et structurées sur les sites. Un des grands principes du Web sémantique est qu'il est nécessaire d'associer aux ressources du Web des informations exploitables par des agents logiciels afin de favoriser l'exploitation de ces ressources. Associer par exemple une notice comprenant des champs : Auteur, Date de création, Date de modification, Mots-clés, à une page Web permet de considérer celle-ci non plus seulement comme comprenant du texte qui ne pourra qu'être traité statistiquement par un robot indexeur, mais également des informations structurées à la sémantique connue et utilisable comme telle par un agent logiciel. De la même manière, si ce qui apparaît comme un simple nombre dans une page Web est de façon explicite marqué comme un couple (valeur, devise), alors un agent pourra faire usage de ces connaissances.

Associer une information exploitable à une ressource signifie deux choses essentielles :

- La première est que cette information doit d'une manière ou d'une autre être structurée – utilisable – et descriptive – de la ressource, de son utilisation – afin de faciliter et d'en améliorer l'accès dans le cas d'une ressource directement visualisée par un utilisateur (par exemple en permettant une recherche d'information plus efficace et plus ciblée), mais aussi l'exploitation dans le cas d'une ressource exploitée dans le cadre d'un service à l'utilisateur (l'utilisateur n'est alors pas forcément conscient de l'utilisation de la ressource). En donnant une sémantique connue aux informations associées aux ressources, il s'agit aussi de leur attribuer un rôle comme par exemple « Auteur ». En effet, la valeur associée à Auteur permettra de rechercher un document dont l'auteur est Jules Vernes, ce qui est nettement plus précis que de rechercher les termes « Auteur » et « Jules Vernes » dans des pages web. Indépendamment du web sémantique et de ces approches et technologies.
- La seconde est que la ressource en question doit exister et pouvoir être exploitée sur le Web indépendamment des informations qui lui sont associées dans le cadre du Web sémantique : celles-ci sont utiles, mais non nécessaires pour accéder et utiliser la ressource, la page Web ou le service⁹.

Le schéma ci-dessous (cf. Figure 29) donne une vision (proposée sur le site <http://www.semanticweb.org/>) de l'utilisation des métadonnées sur le Web sémantique. Des pages Web sont annotées à partir de connaissances disponibles dans une ou plusieurs

⁹ Bien entendu, si la ressource est indépendante de ses méta-données, les méta-données ne sont pas indépendantes de la ressource : si celle-ci évolue, les méta-données peuvent devenir obsolètes voire fausses et inutiles. Mais c'est au créateur des méta-données de gérer la cohérence, et non au gestionnaire de la ressource.

ontologies (qui ont pour objectif de normaliser la sémantique des annotations), et ces annotations, regroupées en entrepôts de métadonnées deviennent utiles pour des agents de recherche d'information, faisant ou non appel à des moteurs d'inférence permettant de déduire de nouvelles connaissances formelles des annotations.

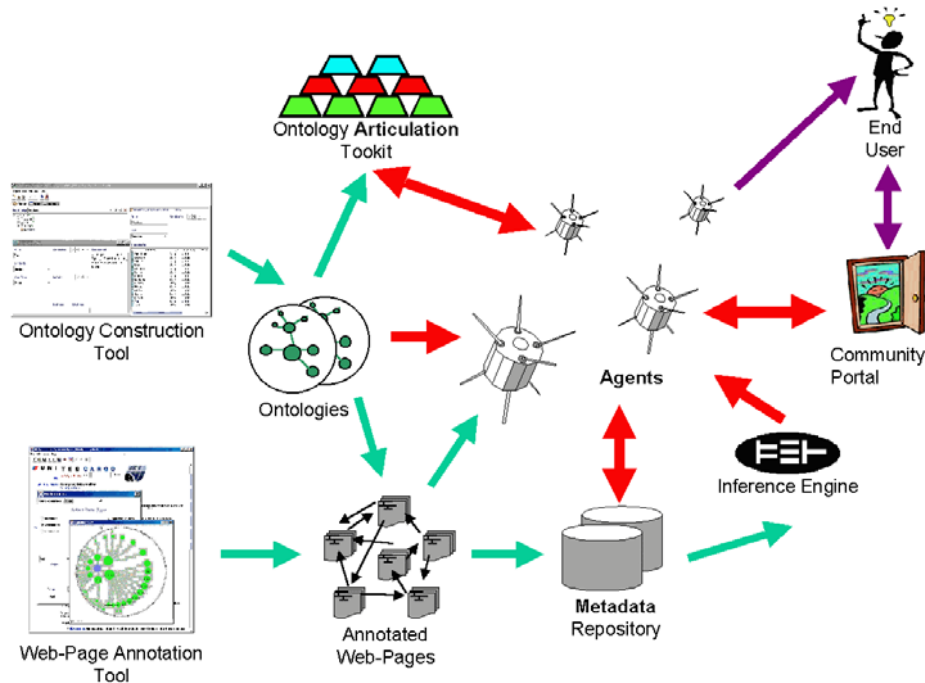


Figure 29 : Utilisation des métadonnées dans le web sémantique

4.4.2.1 Définitions

Au-delà de ce schéma (sur lequel nous reviendrons au travers d'un scénario d'utilisation dans le domaine de l'apprentissage à distance), remarquons que deux termes principaux sont utilisés dans la littérature afin de décrire les informations associées à des ressources : métadonnées et annotations.

Si nous revenons sur ces notions, et de façon générale :

- Une métadonnée est « une donnée sur une donnée ». Cette définition est un peu vague voire ambiguë, et elle est comprise de manière différente par différentes communautés qui conçoivent, créent, décrivent, préservent et utilisent des systèmes d'information et des ressources. Par exemple, dans certains cas la donnée sur laquelle la métadonnée porte est considérée comme ayant le même statut de donnée formalisée, traitable par un système informatique, dans d'autres, la donnée n'est qu'interprétable par un être humain, et seule la métadonnée en permet le traitement automatique.
- Une annotation est à la base une note critique ou explicative accompagnant un texte, et par extension, une quelconque marque de lecture portée sur un document, que celui-ci soit textuel ou image.

On le voit, les termes de métadonnée ou d'annotation prennent bien en compte cette notion d'ajout d'information à une ressource, et on pourra a priori les utiliser indifféremment pour décrire ces informations que le Web sémantique doit ajouter au Web pour le rendre plus utilisable par des machines. Pour autant, si ces deux termes existent, c'est qu'ils n'ont pas le même sens.

La communauté anglophone du Web sémantique, circonspecte sur la différenciation, considère par exemple dans le schéma que nous reprenons plus haut que les annotations de pages Web deviennent des métadonnées dès qu'elles sont stockées dans une base sur un serveur. On différencie alors l'information en tant qu'elle est attachée à (et présentée avec) une ressource et l'information manipulable et exploitable de façon plus indépendante de celle-ci.

D'un point de vue plus lié à la pratique de l'annotation / métadonnées (du point de vue du processus de mise en place), et s'il est nécessaire de distinguer, nous proposons ce distinguo :

- Une métadonnée sera plutôt attachée à une ressource identifiée en tant que telle sur le Web – aura plutôt une pertinence a priori et sera plutôt saisie suivant un schéma. Par exemple, la description normalisée d'un service Web et l'auteur d'un document permettront de mettre en place des inférences.
- Une annotation sera plus située au sein de cette ressource et écrite au cours d'un processus d'annotation / lecture. Par exemple, un commentaire libre associé à un fragment d'une page Web – quelques mots, un paragraphe – déterminé au besoin.

Evidemment, cette distinction n'a rien de définitive, il s'agit simplement de mettre l'accent sur le caractère plus situé au sein de la ressource (du fait de son exploitation par un utilisateur) de l'annotation, par rapport à une métadonnée plus indépendante, voire ressource en tant que telle.

Plusieurs critères, non indépendants – que nous illustrerons dans la suite – peuvent être envisagés pour considérer les différents types de métadonnées (annotations) :

- Les types de ressources qu'elles concernent, plus ou moins fragmentées.
- L'automatisation plus ou moins marquée de leur mise en place.
- La structuration plus ou moins forte de leur « schéma ».
- Les tâches qu'elles soutiennent (ou sous-tendent).
- L'utilisation qui en est faite par les agents logiciels, plus ou moins « intelligente ».
- Le rapport entre coût et gain de ces métadonnées. Plus la description des ressources est poussée plus on pourra en obtenir de « service » ou d'automatisation. Il y a à cela deux limites, l'une est constituée par notre capacité à extraire les connaissances requises et l'autre par le coût pour remplir tous les champs de métadonnées utilisés.

Remarquons que pour le Web sémantique, on parlera le plus souvent de métadonnées ou d'annotations sémantiques. Deux possibilités d'interprétation sont ici disponibles : « sémantique » au sens de « sémantique formelle », et puis au sens plus vague de « faisant sens dans le cadre d'une tâche pour un utilisateur » - encore appelée sémantique interprétative. Si certains tenants du Web sémantique souhaitent limiter la portée de celui-ci au premier sens – formel – ces deux interprétations nous semblent pourtant tout à la fois pertinentes et de toutes façons inextricablement reliées dans le cadre d'un Web sémantique riche et ouvert, à l'image de la richesse et de la diversité du Web actuel, qu'on pourrait qualifier d' « un peu sémantique ».

Dans le cadre projet européen DESIRE (<http://www.desire.org/>), la définition suivante a été proposée pour métadonnées (annotations) :

- Définition : les métadonnées sont des données associées à des objets qui libèrent les usagers potentiels de la nécessité de connaître à l'avance leur existence et leurs caractéristiques.

Cette définition est intéressante car elle met bien en exergue que les métadonnées peuvent être des informations bibliographiques, des résumés, des termes d'indexation, des abstracts, des propriétés spécifiques, tout ce qui peut être un substitut au document original. Ce sont en effet ces métadonnées qui serviront pour retrouver les ressources, les réutiliser, les échanger, les partager et les composer. On accèdera et on utilisera ces ressources qu'au moment de les « consommer ». Les métadonnées doivent donc fournir les informations nécessaires à la réutilisation des informations. Elles déterminent les conditions et les contextes de cette réutilisation. Par des utilisateurs, c'est eux qui assurent l'interprétation de ces données. Par des machines, c'est elles qui assurent l'interprétation et qui automatisent partiellement ou totalement cette réutilisation. Par les deux, certains champs de métadonnées seront dédiés à l'utilisateur ou à la machine. L'utilisation de ces métadonnées va être illustrée dans le paragraphe suivant à l'aide d'un scénario de E-Learning.

4.4.2.2 Scénario d'utilisation des métadonnées et de l'adaptation

Nous présentons dans la suite un scénario d'utilisation du web sémantique articulé autour de l'enseignement à distance. Ce scénario s'intéresse autant à l'utilisation des métadonnées qu'à celle de l'adaptation qui est présentée dans une autre partie.

« Le professeur Bern de l'université de Stuttgart est impliqué dans la création de cours diplômant sur les réseaux à l'intention d'étudiants de toute la Communauté Européenne. Afin de constituer son cours, il va chercher à utiliser toutes sortes de ressources trouvées sur le Web :

- Des ressources « classiques » telles que des articles scientifiques ou des cours numérisés : cours donnés en présence d'élèves et mis à disposition sous différents formats (textes, vidéos, présentations multimédia).
- Des ressources composées d'éléments proposés par différentes universités ou écoles, structurées par un ou plusieurs professeurs pour être accessibles en ligne, qui constituent des parcours cohérents à destination des étudiants. Certaines de ces

ressources sont prévues pour être composées automatiquement par un logiciel à partir d'une spécification afin d'engendrer dynamiquement un cours personnalisé.

Pour cela, le scénario se divise en trois phases : la première est essentiellement centrée sur la recherche d'information principalement interprétable par un humain, la constitution d'un corpus et son indexation pour créer des cours en ligne manuellement, la seconde phase fait appel un système de composition dynamique de cours adaptatifs spécifiés par un auteur et l'utilisation de métadonnées sémantiques interprétables par un logiciel et la troisième phase met en scène les apprenants.

4.4.2.2.1 Recherche d'information et création manuelle de cours

Dans un premier temps, le Pr. Bern utilise un moteur de recherche de type « Google » pour retrouver des ressources intéressantes pour son cours. Pour cela, il commence par une requête simple : « Network » AND « Course », et obtient quelques millions de réponses, dont un grand nombre sont – dès les premières pages – non pertinentes. En effet, le terme « Network » y prend une sémantique bien différente de celle cherchée (réseaux sociaux, réseaux d'excellence, réseaux d'entreprise) tandis que « course » est présent dans de nombreuses pages dans l'expression « of course ». Découragé par la masse de document à analyser, le Pr. Bern décide de changer de méthode et d'utiliser un moteur de recherche fondé sur une approche de type « Web sémantique » utilisant un schéma de métadonnées (Dublin Core – DC). L'outil lui propose un ensemble de descripteurs dont il peut préciser les valeurs ou domaines de valeurs à l'aide de mots clés. Par rapport à la recherche précédente, il peut désormais préciser si les termes recherchés sont dans le Titre du document (network), son Sujet (network) ainsi que sa Date de création. Par contre, le Format (html, pdf, etc.), le Type (course) et le Langage utilisé (anglais ou allemand) sont des contraintes fournies par son profil utilisateur. Il n'a donc pas à les fournir au système. En spécifiant ainsi un rôle et une sémantique pour les différents termes de sa requête, il lui est possible de mieux cibler sa recherche, et d'obtenir des ressources plus pertinentes. Cela n'est évidemment possible que parce que le moteur, lorsqu'il parcourt le Web utilise les métadonnées Dublin Core associées aux ressources. Cependant, avec un schéma de métadonnées comme le Dublin Core, les auteurs des métadonnées, s'ils sont à peu près tous d'accord sur ce que signifie Titre, ne le sont pas forcément en ce qui concerne le Sujet ou le Type d'un document. Il pourra donc y avoir des ambiguïtés sur le rôle et le sens de ces champs.

Le Pr. Bern a quand même réussi à récupérer bon nombre d'éléments potentiels de cours qu'il va lui être possible d'organiser pour produire un cours complet. Comme au cours de toute navigation sur le Web, pour les parties les plus pertinentes de chaque ressource (ensemble de paragraphes, images), il a ajouté une annotation textuelle décrivant la ou les utilisations qu'il prévoit d'en faire plus tard pour faciliter sa tâche ultérieurement.

Conformément à la pratique de son institution, le Pr. Bern ajoute tout d'abord ces documents à la base générale documentaire de l'université. Pour cela, il va être obligé de modifier certains champs de métadonnées. A la différence de la pratique habituelle sur le Web, un thésaurus propre à l'université est utilisé. Celui-ci décrit le vocabulaire à utiliser dans les différents champs, et sert de guide d'utilisation du schéma. De par le référentiel commun constitué par le thésaurus, le Pr. Bern et ses collègues sont à peu près sûrs d'utiliser les mêmes termes pour décrire les documents, et peuvent utiliser la structure de navigation constituée par le thésaurus pour atteindre ceux-ci. Pour toutes les saisies de métadonnées du

DC, le Pr. Bern n'est pas obligé de remplir tous les champs puisque certains sont obtenus à partir de son identifiant personnel et des références de son université tels que « créateur, éditeur, source, date, format ». Il est par contre bien évident qu'il doit remplir lui-même les champs « titre, sujet, couverture, relation, droits, langage et description »

Un des documents ajoutés à la base possède dans sa description Dublin Core un élément qui va être utilisé automatiquement par un logiciel d'alerte, permettant de prévenir trois abonnés de l'université qu'ils peuvent consulter un article les intéressant potentiellement. L'un d'eux va le consulter, y laisse quelques annotations de lecture (il n'est pas du tout d'accord avec une affirmation de la troisième partie). Une discussion s'engage alors lorsque les deux autres utilisateurs ainsi que le Pr Bern (prévenu que le document a été annoté) s'opposent à cet avis avec véhémence. Chacun faisant référence dans son argumentation à d'autres parties du document.

Afin de mettre en place le cours lui-même, le Pr. Bern va réutiliser les documents. Certains sont réutilisables en l'état (par exemple une figure illustrant particulièrement bien une notion), et d'autres ne le sont pas. En effet, sortis de leur contexte de création, ils ne sont pas intégrables directement dans un parcours cohérent, et demandent une fragmentation / modification / adaptation, ainsi qu'une organisation ad hoc.

Si la précision d'utilisation avec le moteur de recherche utilisant les métadonnées est meilleure qu'avec le moteur classique, et s'appuie sur des métadonnées mises en place de façon simple et peu coûteuse en temps, elle n'est suffisante que pour une utilisation par un utilisateur humain, et non par un logiciel de génération automatique de cours. En effet, il est uniquement possible de l'utiliser pour de la recherche d'information qui sera interprétée (et exploitée) par un humain : même si un schéma tel que le Dublin Core définit une structure de métadonnées, l'exploitation de celle-ci par des logiciels est limitée puisque aucune sémantique – interprétable par une machine – n'est associée aux domaines de valeur des différents champs et donc certaines ressources retrouvées pourraient ne pas convenir. Même si les termes sont les bons dans le bon rôle, ils peuvent ne pas avoir été utilisés avec une même sémantique, ou encore la qualité de la saisie des métadonnées n'a pas été validée ou vérifiée.

4.4.2.2.2 Composition dynamique de cours adaptatifs

Pour composer son cours, le professeur Bern utilise un nouvel environnement appelé « E-Learning 2010 ». Ce dernier engendre dynamiquement un cours adapté à chaque étudiant (connu d'après son profil d'étude). « E-Learning 2010 » calcule un cours adapté à chaque étudiant à partir d'une structure de cours choisie par un professeur et d'une spécification du contenu selon un schéma de métadonnées appelé « LOM+ », ainsi que des modes d'évaluation des contenus et de leur utilisation par les méthodes d'adaptation. LOM+ utilise des ontologies comme vocabulaire partagé pour les termes du domaine du cours (ici les « réseaux informatiques »), pour les modèles de cours en fonction d'une approche pédagogique choisie (une structure globale de cours qui organise l'accès aux contenus qui sont typés : présentation d'une notion, puis exercice, approfondissement, etc.), enfin pour les modes d'évaluation et les méthodes d'adaptation pour un modèle utilisateur donné (Garlatti and Prié 2005).

Toutes les ressources qu'il a trouvées / définies précédemment peuvent être réutilisées, mais il est nécessaire d'y associer de nouvelles métadonnées LOM+ pour leur permettre

d'être insérées dans la trame du cours. La saisie des métadonnées à partir de « E-Learning 2010 » est bien plus simple que précédemment puisque des domaines de valeurs – définies par différentes ontologies (domaine, approche pédagogique, modèle utilisateur, adaptation) – sont associées aux différents champs. Pour chaque champ, l'outil propose soit une liste de termes, soit une valeur déduite par le système en fonction du contexte, qu'il suffit de valider ou non selon le type de champ (pour le champ « Status », dans la catégorie « Lyfecycle » la valeur est directement proposée par le système à partir du contexte courant, par contre pour les champs « Difficulty » et « IntendedEnduserRole », les valeurs suivantes sont respectivement proposées : <low, average, high> and <PassiveLearner, ActiveLearner, CollaborativeLearner, TutorLearner, ProfesseurLearner>). Sont fournis également un guide d'utilisation du champ muni d'exemples et de contre-exemples et un navigateur d'ontologie permettant de mieux comprendre la structure de celle-ci et donc la sémantique des concepts ou relations liés à un terme donné. On voit bien dans l'exemple précédent que sans explication ou guide d'utilisation, il aurait été difficile au professeur Bern de fixer une valeur pour ces champs.

Il est possible de rechercher plus précisément de nouvelles ressources qui sont disponibles à partir de « E-Learning 2010 ». Pour cela, il lui suffit de sélectionner un élément du modèle de cours muni d'un type – exercice, solution, problème, définition, approfondissement, explication, éclaircissement, etc. – et une interface de recherche d'information est proposée n'utilisant que les champs de métadonnées pertinents dont certains sont déjà définis comme par exemple le type d'élément. Pour les autres une liste de valeurs est disponible. S'il n'est pas possible au système de trouver dans la base de « E-Learning 2010 », les ressources nécessaires, il lui est possible de réaliser une recherche élargie sur Internet. Cette recherche étant guidée par les ontologies disponibles, il n'y aura plus d'ambiguïté sur les sens des termes utilisés. En effet, il est possible au logiciel de recherche de rechercher toutes les ressources indexées à partir de ces mêmes ontologies ou d'une partie d'entre elles ou encore d'élargir à d'autres ontologies équivalentes mais avec lesquelles il sera nécessaire de réécrire les termes pour assurer une transcription jugée sémantiquement juste. Il lui est également possible de fixer certains critères d'évaluation dans son profil utilisateur pour une recherche adaptative afin de filtrer les résultats ou de le guider dans sa recherche.

Après avoir choisi une approche pédagogique, le professeur Bern sélectionne une structure de cours particulière qui convient à sa manière d'enseigner. Il obtient ainsi une structure globale de cours dans laquelle il doit spécifier des contenus à partir du schéma de métadonnées « LOM+ ». Dans ce schéma de métadonnées, certaines entrées sont dédiées à l'adaptation : par exemple les pré-requis qui prennent leurs valeurs dans l'ontologie du domaine, le type d'étudiants – 1^{re} année, 2^e année, etc. C'est cette spécification qui permettra alors au système de générer les cours en fonction des profils des étudiants - qui ont parmi leurs caractéristiques utilisateur un modèle de recouvrement permettant de connaître pour chaque concept du modèle de domaine le niveau de connaissance acquis ou supposé acquis par l'étudiant. Pour l'adaptation, il doit aussi préciser pour différents stéréotypes d'étudiants les méthodes d'adaptation autorisées, les modes d'évaluation des contenus. Chaque méthode d'adaptation utilise le résultat de l'évaluation des fragments.

Pour la constitution du cours, le professeur Bern a parfois eu des difficultés à bien comprendre la méthodologie de conception liée à « E-Learning 2010 », car cet environnement est assez complexe. Cependant, en plus de la documentation des diverses ontologies et schémas de métadonnées, une assistance logicielle et humaine tant au niveau interne

qu'externe lui est proposée pendant la conception. Sa tâche est à la fois rendue plus complexe sur certains aspects mais aussi grandement facilitée pour d'autres. Ses cours ou éléments de cours sont plus facilement réutilisables pour lui-même et pour les autres puisqu'il lui a fallu avoir une réflexion sur la granularité – taille de chaque élément et niveau d'abstraction pour ceux qui ne sont pas atomiques – des éléments et leur réutilisation. Un suivi de la qualité des métadonnées est mis également en place au sein de l'université afin d'assurer la cohérence de chaque cours ou élément de cours disponible en ligne. Ces dernières doivent être saisies de manière bien plus rigoureuse lorsqu'une partie des champs est destinée à un logiciel de composition de cours. En effet, si l'intention de l'auteur lors du choix d'une valeur de champ n'est pas conforme à l'utilisation que va en faire le logiciel de composition, le résultat ne sera guère prévisible.

4.4.2.2.3 Lecture et apprentissage par un apprenant

Chaque étudiant accédant à « E-Learning 2010 » peut choisir un cours et une approche pédagogique particulière (par résolution de problème, collaborative, etc.). En fonction de son niveau scolaire et de ces acquis précédents, certains cours complémentaires – unités de valeurs – peuvent lui être proposés à partir de son profil utilisateur. En effet, ces unités de valeurs peuvent être indispensables à une bonne compréhension. Dans certains cas, ces différents éléments peuvent aussi résulter d'une négociation avec son tuteur afin par exemple d'assurer une charge de travail raisonnable, ainsi qu'une progression satisfaisante de son apprentissage. Cet étudiant étant à l'université du Sussex, il aura à sa disposition une présentation et une structure de ses pages de cours – contenu, barre de navigation et fonctions – qui sont communes à tous les étudiants du Sussex. En effet ceux de l'université de Stuttgart n'ont pas tout à fait les mêmes habitudes, et ont donc un environnement légèrement différent. Chaque étudiant peut potentiellement disposer d'une structure globale de cours et d'un contenu différent. Néanmoins, chacun peut accéder à l'ensemble s'il le désire.

Chaque étudiant se doit d'être au courant de son propre modèle utilisateur et de ses parties privées / publiques. Un étudiant possède un modèle utilisateur global qui peut être enrichi dynamiquement en fonction des cours auxquels il est inscrit. Pour chaque cours, avant d'accepter l'enrichissement de son modèle, une négociation a été faite entre le système, le tuteur et l'apprenant afin de se mettre d'accord sur les données privées / publiques. En effet, certaines sont obligatoires pour le suivi de l'apprenant par un tuteur et d'autres peuvent être gardées privées ou publiques – pour partager avec d'autres apprenants par exemple. Tout au long de l'apprentissage, ce modèle global est maintenu et enrichi au fur et à mesure qu'il progresse dans ses études. Comme ces informations sont disponibles sur un serveur de modèles utilisateur, il lui est tout à fait possible de continuer ses études dans un autre pays et de continuer à utiliser son modèle. Ceci n'est possible que parce que des vocabulaires communs – ontologies – sont partagés entre différents pays.

On le voit dans cet exemple, si le simple ajout de métadonnées à une ressource (document ou fragment) n'est le plus souvent qu'orienté vers une tâche de « recherche d'information » dans toute sa généralité (schéma général), des schémas de description des ressources peuvent être spécialisés pour des tâches (lecture active) ou des domaines d'application particuliers. Par exemple la construction ou l'utilisation de cours en ligne, les échanges liés à l'information (XMLNews) ou aux produits bancaires, les documents audiovisuels (MPEG7). Une même ressource peut bien entendu être décrite suivant plusieurs espaces d'indexation (points de

vue). Ces schémas de métadonnées se résument le plus souvent en des standards descriptifs exprimés sous la forme de DTD ou de schéma XML qui dans de nombreux cas, n'utilisent bien souvent que des mots clés, voire au mieux des vocabulaires standardisés tels que AAT, ULAN. Ils ne permettent alors ni de définir de manière précise la sémantique opérationnelle (liée à des inférences automatiques) des différents champs des descriptions, ni celles de leurs domaines de valeurs. L'étape suivante proposée par le Web sémantique est donc celle des schémas de métadonnées fondés sur des ontologies, autorisant des inférences formelles sur les métadonnées ou les annotations.

Le scénario que nous venons de présenter illustre un certain nombre de points importants liés aux annotations / métadonnées :

- Il est possible d'annoter une ressource dans son ensemble, ou un fragment de ressource documentaire (pour peu que la ressource soit fragmentable : une image ou un texte l'est, un service de réservation d'hôtel l'est moins immédiatement).
- Les schémas de métadonnées / annotation peuvent être plus ou moins généraux ou spécialisés : du Dublin Core permettant de décrire tout type de documents avec des descripteurs basiques, à une ontologie formelle permettant d'annoter précisément des fragments de cours, ou des dépêches d'agence pour la veille financière (rachat, annonce de perte, etc.), beaucoup de choses sont possibles.
- La mise en place des annotations peut être plus ou moins automatisée : de la mise en place manuelle (« tel paragraphe est intéressant, je vais le réutiliser tel quel »), à l'automatique (cette suite de chiffres suivie du caractère € signifie qu'il convient de l'annoter avec le concept Prix), en passant par le semi-automatique (je vous propose de considérer que ce « 37000 » est un code postal, à vous de me le confirmer).
- Les métadonnées / annotations peuvent être utilisées ou connues des utilisateurs finaux (cas des données permettant de rechercher avec le Dublin Core), ou bien uniquement par des spécialistes (les concepts de description d'une progression de cours ne sont connus que des enseignants).
- Les métadonnées / annotations sont utilisées pour deux grandes tâches. La première est la recherche d'information, puisque toute métadonnée informatique liée à une ressource représente de fait un index pour cette ressource, et peut être utilisée comme telle. La seconde est la composition de documents, de fragments de documents, de services en vue de construire de nouveaux documents (cours à la carte, catalogues, réponses à une question, etc.) ou de nouveaux services (hôtel + spectacle).
- Les schémas de métadonnées sont plus ou moins formalisés et permettent d'offrir des services considérés comme plus ou moins intelligents, puisque les agents logiciels peuvent plus ou moins s'en emparer pour mener des inférences.
- Les annotations et métadonnées peuvent rester dans la sphère personnelle (mes annotations sur une page Web), ou s'inscrire dans une collectivité plus ou moins restreinte (mon groupe de travail, mes étudiants, l'ensemble des usagers du Web). L'annotation peut être mise en place de façon collaborative, être plus ou moins partagée ou publiée. On peut avoir plus ou moins confiance en des métadonnées.

- Point plus technique : les métadonnées / annotations de ressources du Web peuvent être stockées ou mises à disposition dans ou hors de cette ressource : les documents peuvent être « équipés » par leurs métadonnées (alors un moteur pourra collecter celles-ci), ou bien on pourra avoir de véritables bases de (méta-) données centralisant celles-ci. La diffusion des métadonnées peut se faire de façon centralisée ou bien de façon plus directe (P2P).

4.4.3 Conclusion

L'un des principaux problèmes du Web est de trouver l'information ou le service dont l'utilisateur a besoin dans un espace d'information vaste et peu ou pas structuré. Le web sémantique et plus particulièrement les méta descriptions de contenu permettent d'améliorer les méthodes de recherche d'information fondées sur des mots clés, car la sémantique apportée par les métadonnées associée aux ontologies permet de disposer d'un contexte d'interprétation unique des termes utilisés.

L'interopérabilité des systèmes est accrue grâce à l'utilisation d'ontologies partagées. En effet, les données d'un patient figurant dans divers hôpitaux peuvent être combinées pour une meilleure utilisation, les bases de données des musées deviennent compatibles. La gestion de connaissance est devenue nécessaire pour les entreprises désireuses de maintenir, retrouver et réutiliser leur savoir-faire et la connaissance interne.

La conception des systèmes d'information adaptatifs et sémantiques implique une forte transition d'un point de vue conceptuel - notamment sur la ou les méthodologies de conception - et technologique. Nous avons montré comment les documents virtuels, le sémantique web et les métadonnées nous apportent des méthodes et des outils pour la conception de ces systèmes d'information. Ces derniers sont nécessaires mais pas encore suffisants pour fournir des méthodologies de conception. Nous allons maintenant analyser comment nous avons résolu ces problèmes dans un cadre restreint qui nous a emmené à l'environnement SCARCE.

4.5 L'environnement SCARCE

Les systèmes d'information adaptatifs sur le Web ont la capacité à fournir différentes catégories de contenu, de mécanismes de navigation et de présentations en fonction des besoins utilisateur (Brusilovsky 1996). Les systèmes traditionnels ont un modèle de domaine, un modèle utilisateur et des règles d'adaptation. Le modèle de domaine est utilisé pour donner une sémantique aux ressources. Le contenu est généralement connu de l'auteur de l'application et sous son contrôle (Aroyo, De Bra et al. 2003). La structure du site Web correspondant, son contenu, les stratégies d'adaptation et la présentation sont conçus par le même auteur et sont souvent imbriqués. En effet, dans certains systèmes, les règles d'adaptation sont placées avec le contenu dans les pages web.

De nos jours, les systèmes d'information (SI) utilisent partiellement le web comme espace d'information. Ils doivent donc avoir la capacité à réutiliser ces ressources distribués (Brusilovsky and Su 2002; Aroyo, De Bra et al. 2003) et/ou les web services (Conlan, Lewis et al. 2003). On peut voir un SI adaptatif comme un système constitué de différents composants : une organisation, un contenu, des stratégies d'adaptation et une présentation. Ils

peuvent être conçus séparément par différents auteurs géographiquement distribués. Il est donc nécessaire d'avoir une approche méthodologique permettant la réutilisation des contenus. On peut citer quelques enjeux liés à cette problématique : interopérabilité et normes du W3C, la cohérence du document fourni et la séparation des spécifications et de la gestion des composants du SI. En raison de la complexité de ses systèmes d'information, il est important d'avoir des processus de conception et de développement plus rigoureux. Dès qu'un SI est généré dynamiquement à partir de ressources distribuées, la cohérence et la compréhension du système fourni est étroitement liée au contenu, à l'organisation et aux stratégies d'adaptation dédiées aux tâches utilisateur. Il est donc nécessaire de proposer une approche globale unifiant les différents composants et satisfaisant les besoins utilisateur.

Nous nous sommes intéressés aux SI adaptatifs pour lesquels les utilisateurs appartiennent à une communauté de pratiques : ils partagent des connaissances communes pour travailler ensemble (Wenger 1998). Nous proposons d'utiliser les connaissances explicites des communautés pour automatiser la génération d'un SI adaptatif et pour assurer ainsi la cohérence et la compréhension. Les spécifications et la gestion de la sélection, de l'organisation et des stratégies d'adaptation - du filtrage du contenu - sont séparées et fondées sur les connaissances d'une communauté de pratiques. Ainsi, la conception d'un SI adaptatif devient un processus entièrement fondé sur des connaissances. Nous avons conçu un SI adaptatif comme un moteur de composition flexible, appelé SCARCE – SemantiC & Adaptive Retrieval and Composition Environment - fondé sur une approche de type web sémantique dans laquelle la connaissance explicite des communautés de pratiques est formalisée dans des ontologies. Notre approche est constitué: (i) d'une organisation sémantique des ressources, (ii) de spécifications déclaratives pour l'adaptation et (iii) d'un moteur de composition fondé sur des connaissances. Les avantages principaux sont : i) une architecture générique qui est réutilisable dans différents contextes, ii) cette architecture est fondée sur la connaissance explicite des communautés et fournit une méthode pour la conception de SI adaptatifs. En effet, tout nouveau SI peut être produit dès que les spécifications de la sélection, du filtrage – adaptation - ou de l'organisation seront modifiées. Bien évidemment, un tel procédé est limité par les principes fondamentaux sous-tendant le moteur de composition. SCARCE est le cœur des projets ICCARS et CANDLE.

Dans un premier temps, quelques principes de conception des SI adaptatifs sont présentés. Puis, nous développons les principes de conception de notre architecture et notre méthodologie. Finalement, nous présentons brièvement le moteur adaptatif de composition.

4.5.1 Quelques principes pour la conception des systèmes d'information adaptatifs

Pour la conception des SI adaptatifs, la réutilisation des ressources engendre de nouveaux besoins et de nouvelles contraintes. Nous allons analyser brièvement certains d'entre eux : interopérabilité et normes du W3C, cohérence du document fourni, la séparation de la gestion et de la spécification des composants d'un SI.

Interopérabilité et normes : l'interopérabilité est fonction de l'adoption par tous d'une architecture commune et de normes communes pour le sémantique web. Dans l'architecture du sémantique web, l'interopérabilité est assurée à trois niveaux : i) Au niveau adressage, l'URL/URI permet de référencer des ressources quel que soit leur localisation et leur nature.

En effet, des textes, des programmes, des images, des vidéo etc. peuvent être des ressources ; ii) Au niveau logique ou syntaxique, où les ressources peuvent être décrites en XML et avoir une DTD qui définit leur structure logique. Ces ressources XML sont indépendantes de la présentation. En effet, Il est possible de réutiliser ces ressources avec différentes présentations dans différents contextes. Néanmoins, ces ressources XML n'ont pas de sémantique interprétable par des machines. Les humains peuvent interpréter les balises XML, mais cette interprétation peut différer de celle des auteurs. La réutilisation de ces ressources peut se faire dès qu'elles sont conformes à la DTD de référence. Elles peuvent en effet être analysées et transformées à volonté en fonction des besoins. iii) Au niveau sémantique, où les ressources ont des métadonnées déterminant les contextes, les conditions et les règles pour leur recherche et leur réutilisation. Les moteurs de recherche intelligents – fondés sur des moteurs d'inférence - fournissent des outils de recherche d'information en utilisant des métadonnées sémantiques et les ontologies correspondantes associées aux ressources. Au niveau sémantique, la problématique de la réutilisation, du partage et de l'échange des ressources est directement liée à celle des métadonnées et de leurs ontologies.

De plus, trois vues différentes d'un SI adaptatif peuvent coexister dans une telle architecture : sémantique, logique et présentation (Christophides 1998). Chaque vue possède une structure spécifique qui l'organise. La structure sémantique définit l'organisation de la signification du contenu. La structure logique reflète l'organisation syntaxique d'une page web par exemple. La présentation décrit comment les pages web apparaissent sur un terminal. Les vues logique et de présentation sont imbriquées dans un document de HTML, tandis que la norme de XML nous permet de les gérer séparément. Le niveau sémantique est souvent implicite dans les SI adaptatifs traditionnels. Les nouvelles architectures de SI adaptatif, fondées sur le web sémantique, vont permettre de représenter explicitement le niveau sémantique. Il est donc possible de représenter les connaissances explicites de la communauté de pratiques et de les partager. En effet, une telle connaissance peut être une ressource en utilisant un standard, comme OWL (Ontology Web Language).

Pour s'adapter à l'architecture du web sémantique, ces trois vues peuvent être reliées aux trois niveaux du web sémantique comme suit : i) sémantique : logique, ontologie, RDFS/RDF, ii) logique : niveau syntaxique en XML, iii) présentation : XSL/XSLT par exemple pour générer du HTML. Les nouveaux SI adaptatifs peuvent avoir une architecture munie de ces trois vues pour assurer l'interopérabilité. La génération du SI adaptatif pour un utilisateur donné peut donc se décomposer en trois processus de composition séquentiels traitant respectivement les trois vues sémantique, logique et présentation.

4.5.2 Spécifications et gestion séparées des composants des systèmes d'information

Nous analysons maintenant le contenu, la sélection, l'adaptation et l'organisation. Le contenu peut être réutilisé dans différents contextes et différents SI adaptatifs. Il est donc nécessaire de dissocier l'organisation et le contenu et de concevoir un mécanisme de recherche d'information capable de sélectionner le contenu pertinent en fonction de l'utilisateur à partir d'une requête utilisateur ou de spécifications de contenu. Un schéma de métadonnées permet aux concepteurs d'associer des rôles / fonctions aux index et d'améliorer la qualité de la recherche d'information. Avec des métadonnées sémantiques, les index ont des valeurs liées à une ou des ontologies qui déterminent des vocabulaires formalisés et

partagés. Dans notre contexte, le rôle des métadonnées est d'automatiser la réutilisation du contenu, au contraire du schéma de métadonnées de LOM (LTSC 1999) qui n'a pas été conçu pour automatiser la réutilisation des objets d'apprentissage. Il est difficile d'automatiser cette tâche, parce que les auteurs des métadonnées, les ontologies du SI et les utilisateurs doivent partager des connaissances explicites et implicites pour assurer la cohérence de l'ensemble et la qualité des résultats produits. La qualité des métadonnées est cruciale pour rechercher les ressources pertinentes. Il y a toujours un espace d'information implicite correspondant à un schéma de métadonnées : toutes les ressources liées par les métadonnées correspondantes.

De la même manière, les règles d'adaptation ne peuvent pas être incluses ou associées aux ressources puisqu'elles ne sont ni connues à l'avance, ni sous le contrôle des auteurs. Comme la sélection du contenu est faite au niveau sémantique grâce aux métadonnées et leurs ontologies correspondantes, les règles d'adaptation doivent apparier le modèle utilisateur et les métadonnées. Elles doivent être conformes au mécanisme de sélection et aux principes de conception du schéma de métadonnées. Ces principes définissent au moins les propriétés suivantes : i) le niveau de granularité pour la recherche d'information et pour la réutilisation ; ii) la répartition du rôle des index pour la sélection et l'adaptation. Pour le niveau de granularité, les ressources sélectionnées peuvent être réutilisées complètement ou partiellement. Dans le schéma de métadonnées, deux sous-ensembles d'index différents ou non peuvent être utilisés : l'un pour la sélection des ressources pertinentes et l'autre pour l'adaptation qui n'est qu'un raffinement de la sélection. Les règles d'adaptation et les requêtes pour la sélection des ressources doivent prendre en compte ces principes.

Dans les SI adaptatifs, l'organisation est généralement un graphe orienté permettant au concepteur de fournir un accès au contenu par navigation à l'utilisateur. Ce graphe peut être explicitement représenté ou être calculé dynamiquement. Ce type d'organisation peut être défini au niveau sémantique à l'aide d'ontologies, et ainsi être étroitement lié aux connaissances explicites de la communauté de pratiques. Le contenu, l'adaptation et l'organisation doivent faire face à différents besoins. Mais, ils sont également confrontés à un but commun : comment assurer la cohérence du SI produit ? La cohérence et la compréhension d'un SI adaptatif sont étroitement liées au contenu, à l'organisation, à la présentation et aux stratégies d'adaptation dédiées aux tâches utilisateur. Dès qu'un SI adaptatif est généré dynamiquement à partir de ressources distribués, il est plus difficile d'assurer la cohérence du document fourni. Ainsi, il est nécessaire d'avoir une méthodologie qui fournit une approche globale unifiant les différents composants.

4.5.3 Principes de conception

L'automatisation des SI adaptatifs et de la réutilisation des ressources peuvent être réalisées par des documents virtuels (DV). Les DV sont des documents pour lesquels le contenu, l'organisation et la présentation sont créés à la demande (Milosavljevic, Vitali et al. 1999). Dans notre approche, nous considérons un SI adaptatif comme un DV adaptatif fondé sur des ontologies et défini comme suit :

- Un DV adaptatif se compose d'un ensemble de ressources, de leurs métadonnées et les ontologies correspondantes et d'un moteur de composition qui peut sélectionner et filtrer les ressources pertinentes - adaptation, les organiser et les assembler en adaptant différents aspects visibles du document délivré à l'utilisateur.

Notre moteur de composition est constitué de trois moteurs de composition : sémantique, logique et présentation. La sélection, le filtrage et l'organisation sont traités au niveau sémantique et l'assemblage aux niveaux logique et présentation. Les ressources réutilisables munies de métadonnées sont appelées des fragments. Nous considérons deux types de fragments : les fragments atomiques qui sont des unités d'information et ne peuvent pas être décomposés ; les fragments abstraits qui se composent d'autres fragments (atomiques ou abstraits) et d'une ou plusieurs structures les organisant. Nous avons un modèle de document et un modèle d'adaptation pour lier le contenu, l'adaptation et l'organisation de manière cohérente. Ces modèles sont instanciés par un auteur sous la forme d'un document générique qui spécifie l'organisation, la sélection et l'adaptation au niveau sémantique. Le document générique est « étendu » et instancié à l'exécution pour générer dynamiquement un graphe sémantique. C'est une représentation sémantique du SI adapté aux besoins utilisateur.

En résumé, nous avons besoin d'un modèle utilisateur et d'un modèle de domaine, un schéma de métadonnées et un modèle de document. Ces modèles sont formalisés dans différentes ontologies : l'ontologie des métadonnées au niveau information qui décrit la structure d'indexation des ressources, quelques valeurs d'index sont prises dans les ontologies de domaine et de document ; l'ontologie de domaine représentant la connaissance dans un domaine spécifique ; l'ontologie de document se composant d'un modèle de document et d'un modèle d'adaptation ; l'ontologie utilisateur qui définit différents stéréotypes et différentes caractéristiques individuelles.

Pour commencer, nous présentons les principes fondamentaux de l'organisation, de la sélection et du filtrage ou adaptation.

- L'organisation et la sélection sont associées. L'organisation est fondée sur un graphe générique orienté muni de nœuds génériques, d'arcs génériques et d'une racine unique. Les nœuds génériques possèdent des spécifications de contenu et les arcs génériques sont des relations sémantiques. Les propriétés du graphe générique sont décrites dans l'ontologie de document.
- La sélection est un processus de recherche d'information sur un ensemble de ressources (locales ou distribuées) indexées par un schéma de métadonnées. Les métadonnées sont utilisées pour la recherche d'information et le filtrage qui est un raffinement de la sélection. Un sous-ensemble du schéma de métadonnées est dédié aux spécifications du contenu et un autre sous-ensemble pour les spécifications de l'adaptation. Le même niveau de granularité (le fragment) est employé pour la sélection et la réutilisation.
- Le filtrage ou adaptation permet de gérer des méthodes de navigation adaptative. Le principe est le suivant :
 - Dans un premier temps, les ressources sélectionnées sont évaluées : le but de l'évaluation est de placer chaque ressource dans une classe d'équivalence en fonction des règles d'appartenance à la classe. Ces classes sont déclarées dans l'ontologie de document en fonction des connaissances de la communauté de pratiques. Un maximum de cinq classes d'équivalence peut être défini. Il serait en effet difficile pour un utilisateur de traiter trop de classes d'équivalence (Brusilovsky 1996; Bra 1999).

- Dans un second temps, une méthode adaptative est choisie pour l'utilisateur courant, à partir de ses préférences et des stéréotypes associés aux méthodes de navigation adaptative. Cette méthode est appliquée pour masquer, annoter, trier, etc. les classes d'équivalence de ressources. Si les ressources d'une classe d'équivalence sont masquées à l'utilisateur, il y a aussi adaptation du contenu puisque celui-ci n'est plus accessible à l'utilisateur.

4.5.3.1 L'ontologie de document

Elle se compose d'un modèle de document et d'un modèle d'adaptation. Ces deux modèles constituent le cœur de SCARCE. Le modèle de document est fondé sur le concept de document générique qui se compose de fragments atomiques et abstraits (cf. Figure 30). Ces fragments sont spécialisés en sous-catégories. Les fragments abstraits sont notamment constitués de relations sémantiques. Les sous-catégories et les relations sémantiques dépendent de la communauté de pratiques et sont donc modifiées en fonction de l'application envisagée. Elles font partie des connaissances communes partagées par les membres de la communauté de pratiques.

Dans les projets CANDLE et ICCARS, les auteurs ont un savoir-faire qui leur permet de choisir le contenu des documents et de les organiser de manière cohérente avec une ou plusieurs stratégies de lecture. Le contenu et l'organisation sont sémantiquement reliés pour assurer la compréhension du lecteur. Nous appelons ces organisations des structures narratives. Le lecteur a la capacité d'identifier - parfois inconsciemment - ces structures. Par exemple, les articles scientifiques, les polycopiés de cours, les dossiers thématiques en journalisme, etc., ont tous des structures narratives distinctes. Actuellement, la structure narrative est implicite dans un document imprimé, mais également dans un document numérique. Nous faisons ici l'hypothèse que le savoir-faire et les compétences d'un auteur lui permettent de concevoir ces organisations en fonction des besoins des lecteurs, par exemple pour les journalistes et les enseignants. Ce savoir-faire peut être représenté à un niveau connaissance et puis être réutilisé, partagé et échangé avec d'autres auteurs, utilisé pour produire dynamiquement des SI et pour améliorer la compréhension du lecteur.

Dans ces deux projets, les relations sémantiques utilisées sont des relations de la théorie des structures rhétoriques, « Rhetorical Structure Theory » (Mann and Thompson 1988). La RST a été initialement utilisée pour la génération automatique de texte. Elle propose un cadre d'analyse de la structure organisationnelle et fonctionnelle d'un texte en établissant les relations rhétoriques entre des segments (le plus souvent adjacents). Mann et Thompson ont proposé une vingtaine de relations. La définition de ces relations repose sur deux arguments : Le segment noyau et le segment satellite. Chaque relation est définie par des contraintes sur chacun de ses arguments et se traduit par un effet sur le lecteur. Prenons l'exemple de la relation "évidence", l'effet sur le lecteur est d'accentuer son acceptation des propos tenus dans le segment source de la relation. L'intérêt d'un cadre formel tel que la RST pour l'expression des relations rhétoriques repose sur l'utilisation d'un vocabulaire commun, dont la sémantique est unique. Elle permet donc aux auteurs d'explicitier et de justifier l'organisation des dossiers thématiques, par la même elle permet d'explicitier le savoir-faire des auteurs. Cette formalisation améliore la compréhension de la structure narrative et par conséquent le partage et la réutilisation par d'autres auteurs. Si on analyse par exemple tous les dossiers thématiques dédiés aux naufrages de pétroliers sur les côtes bretonnes, on s'aperçoit qu'ils ont

pratiquement la même structure narrative. On pourrait donc avoir des gabarits de dossiers thématiques qui seraient réutilisables avec ou sans modification d'un naufrage à l'autre. Dans le paragraphe dédié à l'acquisition des connaissances pour l'ontologie de domaine, une analyse théorique plus approfondie des apports de la théorie des structures rhétoriques sera présentée pour les deux projets CANDLE et ICCARS.

Dans ICCARS, les sous-catégories de fragments sont les catégories d'articles et de dossiers thématiques produites par les journalistes. Dans CANDLE, il s'agit de toutes les catégories de cours et d'éléments de cours.

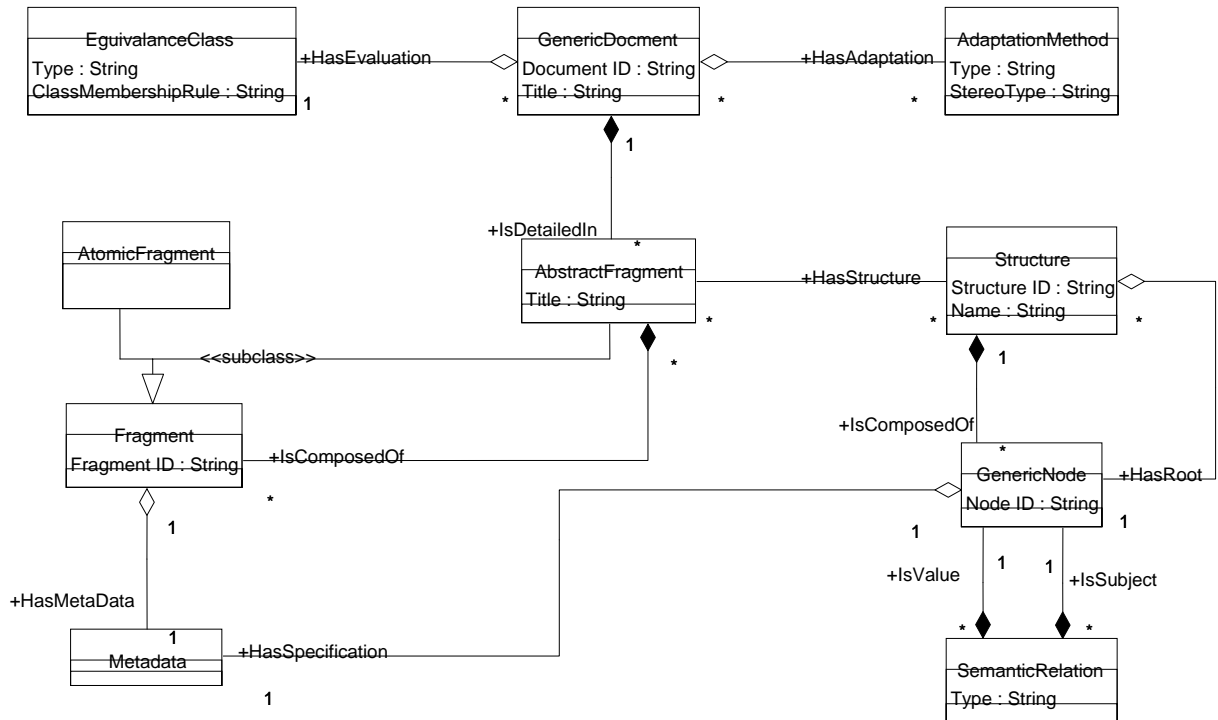


Figure 30 : Modèle de document

Un fragment abstrait peut être organisé selon une ou plusieurs structures. A un fragment abstrait peut donc être associé une ou plusieurs stratégies de lecture. Dans le cas d'ICCARS, un journaliste peut produire plusieurs structures narratives pour un dossier thématique particulier. L'idée est de proposer aux lecteurs plusieurs manières d'aborder le dossier et ainsi de véhiculer plusieurs messages aux lecteurs. Dans le cas de CANDLE, chaque structure narrative est interprétée comme une approche pédagogique différente pour assimiler un même cours.

Une structure est une collection de nœuds génériques parmi lesquels la racine en est un. Un nœud générique est un objet abstrait, qui existe seulement à l'intérieur d'une structure. Il est lié à d'autres par une relation sémantique. Comme nous l'avons vu précédemment, ces relations sémantiques sont spécifiques à chaque communauté de pratiques. L'ensemble composé des nœuds génériques et des relations correspondantes forment l'organisation du SI adaptatif. Un fragment abstrait est un graphe orienté dans lequel les nœuds sont des nœuds génériques et les arcs sont des relations sémantiques entre nœuds génériques. A l'exécution,

un nœud générique peut être vue comme un processus de recherche d'information qui utilise une spécification de contenu faite par l'auteur en fonction des métadonnées, pour faire une requête au moteur de recherche d'information intelligent – moteur d'inférence + ontologie + métadonnées. A chaque spécification de contenu est normalement associée après cette phase de sélection de contenu un faible nombre de ressources. Dans ICCARS et CANDLE, nous avons donc choisi d'associer un espace particulier de l'information à un fragment abstrait pour assurer la cohérence du SI. Il s'agit de limiter le nombre de fragments qui seront sélectionnés par chaque nœud générique afin de s'assurer que l'organisation et le contenu soient cohérents. Néanmoins, cette contrainte peut être relaxée et tous les fragments classés par le schéma de métadonnées pourraient être utilisés. Ils dépendent des communautés de pratiques. C'est un des principes fondamentaux du moteur de composition qui a des conséquences importantes sur la génération du document réel. Des documents génériques peuvent être organisés dans une hiérarchie de généralisation/spécialisation dans laquelle les classes sont des modèles des documents génériques.

Le modèle d'adaptation se compose de classes d'équivalence, des règles d'appartenance aux classes, d'un ensemble de méthodes d'adaptation munies de stéréotypes. Un document générique se compose également de règles d'adaptation qui sont des instances des différentes entités du modèle d'adaptation. Une règle d'appartenance de classe est une comparaison entre les caractéristiques de l'utilisateur (ontologie utilisateur), et les index du fragment (ontologie des métadonnées). Une méthode d'adaptation est autorisée pour un utilisateur lorsque son modèle s'apparie au stéréotype associé à cette méthode.

4.5.3.2 L'ontologie des métadonnées

Le schéma de métadonnées (cf. Tableau 1) décrit la structure des métadonnées et les différents descripteurs des fragments. Ce schéma se compose de deux types de caractéristiques, des descripteurs généralement réutilisables pour des ressources Web telles que l'auteur, la langue et la date de la création, et des descripteurs plus spécifiques qui dépendent de la communauté de pratiques. Les descripteurs réutilisables sont habituellement utilisés dans de nombreux schémas de métadonnées. Ils ne sont pas obligatoires, cela dépend de l'application. Dans le Tableau 1, le fond gris désigne les entrées spécifiques.

Tableau 1. Le schéma de métadonnées d'ICCARS

MD.1	Généralités	Informations générales au sujet de la ressource	
MD.2	Cycle de vie	Entrées pour le versioning	
MD.3	Méta Méta données	Informations sur des méta-données	
MD.4	Technique	Informations techniques sur la ressource	
MD.4.1	Endroit	Où la ressource peut être trouvée ?	
MD.4.2	Format	Format de la ressource	
MD.4.2.1	Type	Type de la ressource (ppt, Doc., html,...)	
MD.4.2.2	Taille	Taille de la ressource en K bytes	
MD.5	Classification	Données au sujet de contenu et des dispositifs de reportage	
MD.5.1	Domaine	La description s'est reliée au domaine	
MD.5.1.1	Concept	Nom de concept	Ontologie de domaine
MD.5.1.2	Niveau	Niveau de la connaissance exigé	
MD.5.2	Reportage	Reportage de la classification	
MD.5.2.1	Type de ressource	Type de ressource (entrevue, rapport...)	Ontologie de document
MD.5.2.2	Édition	Édition concernée	Ontologie de document
MD.5.2.3	Ville	Ville concernée	
MD.6	Droits	Employez les états de la ressource	

Le schéma se compose de six sections, les sections non spécifiques sont habituellement proposées dans tous les schémas de métadonnées (" général », « cycle de vie », « technique »,...). Certaines sections sont plus spécifiques à notre architecture. La section « cycle de vie » possède des descripteurs particuliers pour le traitement des versions (Iksal and Garlatti 2001). La section « classification » doit décrire le contenu du fragment à l'aide du modèle de domaine. Les entrées MD 5.2 définissent le type d'articles écrits par les journalistes. Il dépend bien évidemment de la communauté de pratiques. Il s'agit ici que celle des journalistes impliqués dans le projet : ce sont toutes les catégories et sous-catégories d'articles. Dans ICCARS et CANDLE, les schémas de métadonnées ont de nombreux éléments communs. Par contre celui de CANDLE est proche de celui de LOM avec des extensions.

4.5.3.3 L'ontologie utilisateur

Cette ontologie décrit les caractéristiques utilisateur pour l'adaptation (cf. Tableau 2). Elle est utilisée par le modèle d'adaptation pour définir les règles d'appartenance des classes aux stéréotypes associés aux méthodes adaptatives. L'ontologie utilisateur a besoin de l'ontologie de domaine pour décrire la connaissance de l'utilisateur à l'aide d'un modèle de recouvrement. Généralement, le modèle utilisateur comprend les cinq sections suivantes : personnel, préférences, connaissance, historique et session. Il est possible d'ajouter quelques sections spécifiques selon la communauté de pratiques. La section « Personnel » se compose de données sur l'utilisateur tel que son nom. La section « préférences » est utile pour la personnalisation ; c'est une section qui peut être modifiée par l'utilisateur. La section « connaissance » décrit le modèle de recouvrement et est utilisée par le système pour évaluer

la pertinence des fragments. La section « historique » est employée pour proposer des services à l'utilisateur tels que des signets. La section « session » est uniquement utilisée par le système pour mémoriser la session courante. Les sections « histoire » et « session » sont spécifiques à notre architecture parce qu'elles sont utilisées par le moteur de composition. Même si nous n'avons traité les aspects adaptations dans CANDLE, nous avons réutilisé un modèle utilisateur très proche de celui-ci.

Tableau 2 : Le modèle d'utilisateur d'ICCARS

UM.1	Données personnelles	Données personnelles au sujet de l'utilisateur	
UM.1.1	Identité	Son identité	
UM.1.2	Ouverture	Données de identification uniques	
UM.1.3	Classification	Données de classification	
UM.1.3.1	Endroit	Où habite-t-il ?	
UM.1.3.2	Activité professionnelle	Quel genre de travail ? (économiste, pêcheur, étudiant, etc.)	
UM.1.3.3	Rôle	Le rôle dans l'application (auteur, lecteur...)	
UM.2	Préférences	Préférence de l'utilisateur	
UM.2.1	Intérêt	Matières d'intérêts	
UM.2.1.2	Matière	Une liste de matières	ontologie de domaine
UM.2.2	Adaptation	Préférences d'adaptation	
UM.2.2.1	Élément	Un élément qui peut être adapté (le lien...)	
UM.2.2.2	Règle	Une méthode d'adaptation (annotation...)	ontologie de document
UM.3	Connaissances	Connaissance de l'utilisateur	
UM.3.1	Domaine	La connaissance au sujet du domaine	
UM.3.1.1	Élément	Un concept de domaine	ontologie de domaine
UM.3.1.2	Niveau	Un niveau de la connaissance	
UM.4	Historique	Données au sujet de l'accès aux documents génériques	
UM.5	Session	Données au sujet de la session courante	

4.5.3.4 L'ontologie de domaine

Cette ontologie est un modèle de domaine qui se compose des concepts et des relations du domaine d'application. Elle permet de relier les connaissances utilisateur, la description des fragments et les spécifications de contenu d'un nœud générique. Dans le cas de CANDLE, l'ontologie du domaine est celle des réseaux et des télécommunications. Pour ICCARS, l'ontologie de domaine est celle de la pêche et de la course au large.

4.5.3.5 Acquisition des connaissances pour l'ontologie de document

Dans chaque projet, le processus d'acquisition de connaissances s'est déroulé de manière différente. Nous allons ici présenter ces processus pour les projets ICCARS et CANDLE, ainsi que les fondements théoriques qui ont permis de proposer des modèles de documents adéquats à l'application envisagée.

4.5.3.5.1.1 ICCARS

Pour ICCARS, l'acquisition de connaissance des journalistes et l'évaluation de l'ergonomie du poste de journaliste assisté par ordinateur se sont déroulées dans le cadre du

projet ERGOWEB avec nos collègues du CRPCC de Rennes avec qui nous avons menés les expérimentations nécessaires et sous leur responsabilité d'un point de vue théorique. Nous allons maintenant présenter brièvement le cadre théorique qui leur a permis de réaliser ces expérimentations.

En psychologie cognitive et psycholinguistique, des recherches ont montré que la compréhension des textes est déterminée par la structure globale du texte aussi bien que par la connaissance antérieure du lecteur. Ainsi, quand les scripts ou les schémas (Minski 1975; Fayol and Monteil 1988; Schank and Abelson 1997) sont conformes à la structure rhétorique d'un texte, les expériences ont montré de meilleures performances pour des tâches de rappel et de résumé (Mathews 1982; Englert, Stewart et al. 1988; Ehrlich, Charles et al. 1992). En outre, les résultats prouvent que la structure de rappel du texte préserve généralement la structure rhétorique originale. Ces travaux ont été étendus au domaine de la compréhension de documents multiples (Rouet, Favard et al. 1997; Perfetti, Rouet et al. 1999), où il a été montré que les lecteurs doivent établir les relations entre les documents afin d'élaborer un modèle de situation et que les experts aident les lecteurs à relier des sources d'informations en fournissant ce modèle de situation. On peut donc en conclure que la structure rhétorique d'un dossier thématique - fournissant les relations entre les articles - est importante pour améliorer la compréhension du lecteur. Afin de modéliser la structure globale d'un dossier thématique à l'aide d'une ontologie, il est nécessaire d'explicitier la structure rhétorique des dossiers thématiques. En faisant l'hypothèse que la structure rhétorique des dossiers thématiques est une bonne représentation des scripts et des schémas des journalistes, nous devons donc acquérir les compétences et le savoir-faire des journalistes.

Dans les trois dernières décennies, la psychologie cognitive et la psycholinguistique ont accumulé de très nombreuses données sur la façon dont les êtres humains produisent et comprennent la langue écrite (Fayol 1985; Denhière and Baudet 1992; Coirier, Gaonac'h et al. 1996). Par exemple, la recherche dans ces secteurs a montré que quand quelqu'un lit ou écoute un long texte ou un texte multiple, la compréhension est dépendante de l'organisation globale du texte. Ainsi, certains types de textes (par exemple, les romans, les articles ou les rapports scientifiques) ont une structure conventionnelle qui est bien connue du lecteur. On a montré que la connaissance de ces conventions aide autant le lecteur à comprendre le texte qu'à reproduire plus tard ces mêmes textes (Calfee and Chambliss 1987; Folman and Gissi 1990). Ainsi, certains lecteurs ou auteurs ont des « compétences textuelles » qui leur permettent d'identifier ou de produire de « bons » textes par application de certaines règles. Une telle compétence correspond à un « schéma mental » appelé « script » ou « schéma » (Minski 1975; Fayol and Monteil 1988; Schank and Abelson 1997). Par exemple, Van Dijk et Kintsch (Dijk and Kintsch 1983) ont montré que les articles de journaux ont la structure suivante :

- Article = {sommaire + rapport} ; Sommaire = {titre + introduction topique}.
- Rapport = {l'information + opinion} ; L'information = {l'information précédente + l'information actuelle}.
- L'information actuelle = {situation générale + situation réelle}.
- Situation actuelle = {histoire + événement courant}.
- Histoire = {contexte historique + événements précédents}.

- Événement courant = {contexte + événements réels} ; Événements = {événements + conséquences principaux}.
- Opinion = {commentaire + conclusion}.

Naturellement, ces schémas formels spécifiques sont des idéalizations et ne reflètent pas ce qui se produit réellement dans chaque cas d'écriture ou de lecture. Dans le cadre d'ICCARS, notre but est d'accéder aux schémas activés par les journalistes quand ils produisent des dossiers thématiques. Puis, nous modéliserons les structures rhétoriques utilisées dans ces dossiers par des ontologies.

Les compétences et les savoir-faire d'un journaliste - schémas - ont une représentation implicite dans un dossier thématique numérique. Cette représentation - qui fournit les relations entre articles - peut être analysée par élicitation des connaissances. Dans notre cas, la structure globale d'un dossier thématique se compose des relations et des concepts qui sont des types d'article ou des types de dossier thématique. Ces relations et ces différents types doivent être acquis. Une méthode d'élicitation des connaissances fondée sur des protocoles verbaux est employée avec quatre journalistes. Elle est constituée des trois étapes suivantes :

1. Une entrevue semi-dirigée, fondée sur la question initiale « Pourriez-vous expliquer en détail ce qu'est votre travail quand vous devez réaliser un dossier thématique sur un sujet donné ? », suivi des questions « comment ? » et « pourquoi ? ». Cette première étape vise à obtenir des informations sur les buts, les sous-buts et les procédures réellement suivis par les journalistes produisant un dossier thématique.
2. La production d'un dossier thématique de manière « traditionnelle » (le but est de publier le dossier thématique dans un journal, à l'aide des outils habituels), avec des explications du journaliste (nommant particulièrement le type d'articles et montrant les relations entre les articles composant le dossier thématique). Cette deuxième étape permet d'obtenir des informations sur des schémas activés par le journaliste et la structure rhétorique sous-jacente des dossiers thématiques.
3. La production d'un dossier thématique avec « ICCARS » (le but est de publier un dossier thématique sur le Web, en utilisant le poste de journaliste assisté par ordinateur conçu pour ICCARS), en situation de formation. Des protocoles verbaux sont recueillis à partir des questions des journalistes au formateur. Cette troisième étape permet d'obtenir des informations sur la logique fonctionnelle de l'application et la qualité de l'interface.

Cette méthode nous a permis d'acquérir l'ontologie des dossiers thématiques, c'est-à-dire pour étudier et modéliser les schémas existants chez les journalistes et leurs structures rhétoriques pour des dossiers thématiques afin de concevoir un poste de journaliste assisté par ordinateur adapté à leurs besoins. Nous avons pu ainsi acquérir les sous-catégories de fragments et les modéliser dans une ontologie, mais aussi déterminer la pertinence de la théorie des structures rhétoriques pour définir les relations sémantiques entre les articles dans un dossier thématique.

Après cette expérimentation, nous avons pu finaliser la conception du poste de journaliste assisté par ordinateur qui a pris la forme suivante :

Le poste de JAO possède quatre fonctions principales :

- Visualisation d'articles.
- Composition de dossiers thématiques de presse.
- Recherche d'articles et de sous-dossiers.
- Indexation des articles et des sous-dossiers à l'aide du schéma de métadonnées.

La figure 31 ci dessous montre l'interface du poste de JAO :

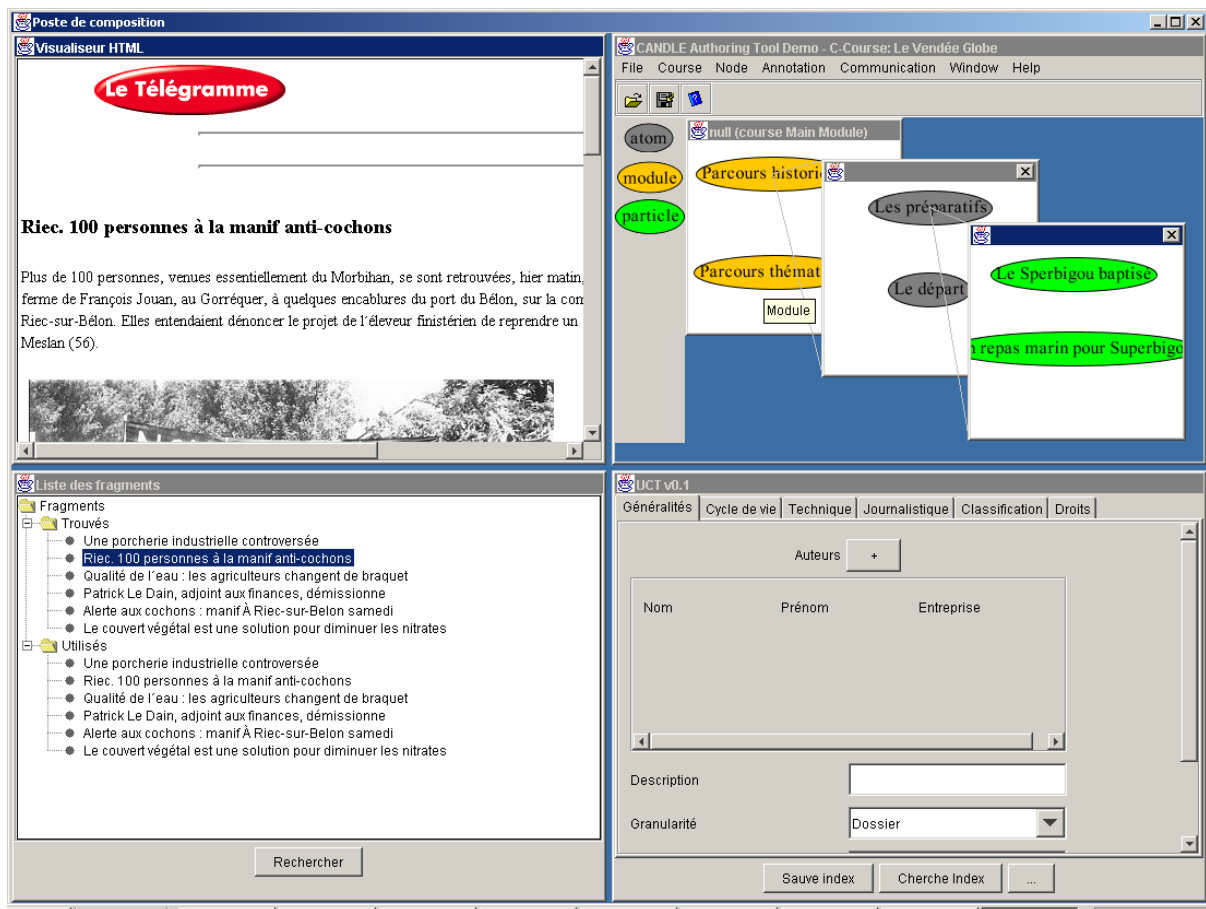


Figure 31 : Poste de journalisme assisté par ordinateur.

4.5.3.5.1.2 CANDLE

Dans le cadre de CANDLE, les cours en ligne sont assez proches de « polycopiés virtuels » engendrés dynamiquement sur des pages web à partir d'une spécification donnée par des enseignants. Toutes les caractéristiques présentées précédemment à propos de la structure des documents, leur structure rhétorique et la facilité de compréhension pour le lecteur peuvent s'appliquer. Nous avons donc réutilisé la théorie des structures rhétoriques

pour modéliser un cours en ligne de ce type. Il y a malgré tout une différence fondamentale avec ICCARS, c'est que les auteurs des cours, impliqués dans le projet n'avait pas un savoir-faire suffisant pour que l'on puisse acquérir leurs connaissances et les formaliser. C'est donc un expert en éducation (Ian Stevenson) qui a proposé un modèle pédagogique fondé sur la théorie des structures rhétoriques et sur la théorie de l'activité. Ce modèle pédagogique a ensuite été utilisé pour « construire » / « créer » la communauté de pratiques qui s'est approprié ce modèle et l'a fait évoluer. Nous allons maintenant présenter brièvement ce modèle à partir d'un livrable du projet CANDLE rédigé par Ian Stevenson (Stevenson 2003).

Le modèle pédagogique est fondé sur la théorie de l'activité et la théorie des structures rhétoriques. Nous allons commencer par une brève définition de la pédagogie qui la relie au contexte, aux buts et aux rôles. Puis, les différents aspects de la théorie de l'activité et des structures rhétoriques sont alors décrits. En conclusion, les six dimensions qui composent le modèle sont présentées, ainsi que des informations détaillées sur chacun de ces éléments.

La pédagogie peut être interprétée comme les interactions qui ont lieu entre les enseignants et les étudiants et les facteurs qui façonnent et orientent ces interactions. La pédagogie doit considérer le contexte et la nature des relations humaines qui structurent l'enseignement et la formation, notamment si c'est à distance ou en présentiel. Une assertion fondamentale de cette approche est que l'apprentissage ne peut se faire que dans l'interaction du contexte, du but, des séquences d'actions et de la rhétorique. Pour prendre en compte ces paramètres, il est nécessaire de faire appel à deux théories : la théorie de l'activité et la théorie des structures rhétoriques. La théorie de l'activité identifie les éléments essentiels de ces systèmes, les utilisateurs et leurs buts et les contextes, y compris la communauté dans laquelle les activités se produisent. Les interrelations entre tous les éléments d'une activité est un point clé de la description des systèmes complexes du comportement tels que la pédagogie. La théorie de l'activité possède deux éléments importants pour cette analyse : système d'activité et structure d'activité. La théorie des structures rhétoriques va permettre d'explicitier comment sont structurées les activités pédagogiques en fonction des modes spécifiques d'interaction entre les enseignants et les étudiants.

Les effets entre le contexte, le but et l'action peuvent être décrits par la notion de système d'activité.

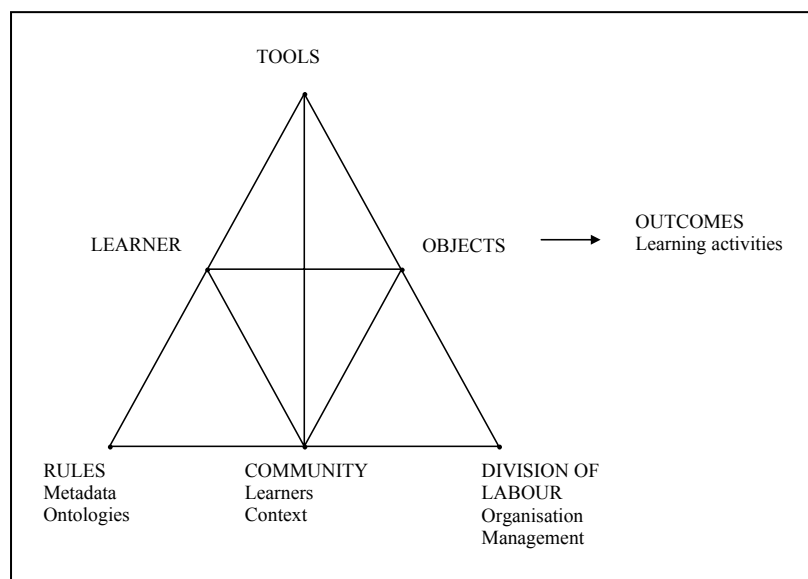


Figure 32 : Système d'activité

La figure 32 représente la structure d'un système d'activité. Le sujet se réfère à l'individu ou au sous-groupe appartenant à l'organisme / l'institution choisi pour l'analyse. L'objet se rapporte à la « matière première » ou à « l'espace de problème » pour lesquelles l'activité est dédiée et qui est façonnée et transformée en résultats à l'aide des outils physiques et symboliques, externes et internes de médiation, y compris des signes. La communauté est constituée d'individus divers et multiples et/ou de sous-groupes qui partagent un même objet général. La division du travail se rapporte à la division horizontale des tâches entre les membres de la communauté et à la division verticale hiérarchique. Enfin les règles se rapportent aux règlements, aux normes et aux conventions explicites et implicites qui contraignent les actions et les interactions dans le système d'activité (<http://www.edu.helsinki.fi/activity/6a0.htm>)

Un deuxième élément de la théorie de l'activité qui est utile ici est la décomposition des activités en niveaux :

- Activité : activité globale aux objectifs bien identifiés.
- Actions : décrit un processus discret avec un objectif clair. Prises ensemble, les actions ont une séquence et une structure qui caractérise l'activité globale.
- Opérations : conditions qui permettent aux actions d'être réalisées.

En identifiant les activités en termes de résultats à atteindre, il est possible d'identifier les conditions et la séquence des actions nécessaires pour atteindre ces résultats.

La théorie des structures rhétoriques (RST) (Mann and Thompson 1988) traite du choix des formes particulières d'expression qui réalise les buts du discours. Elle a été principalement appliquée au texte pour analyser et expliciter les relations entre des phrases, des paragraphes ou encore des blocs de texte. En analyse de discours, la représentation de cette structure du contenu est gérée par l'identification de prédicats rhétoriques - relations. Pour les textes écrits

le choix du prédicat rhétorique nous indique avec précision le type du document et l'intention de l'auteur parce que les prédicats rhétoriques sont utilisés pour développer une argumentation liant ensemble les propositions du contenu - simples phrases par exemple. Les prédicats rhétoriques incluent notamment la cause, conséquence, condition, le but, instrument, similitude, compare et contraste, problème et solution, négation, alternative et collection. Dans le contexte de la pédagogie, la RST peut fournir des mécanismes pour analyser la structure sémantique globale d'une activité, qui accompagne son ordonnancement temporel.

Voici quelques exemples de structures sémantiques pour les activités pédagogiques suivantes :

- Le modèle du cours magistral ou la conférence (cf Figure 33).

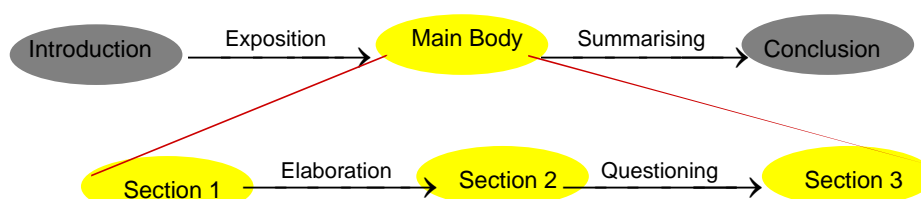


Figure 33 : Structure d'un cours magistral ou d'une conférence

- Le modèle dit du « conflit » (cf. Figure 34) : la résolution des problèmes ou les activités fondées sur la résolution de problèmes dans lesquelles l'enseignant pose un problème pour que les étudiants le résolvent. L'enseignant demande alors aux étudiants de trouver une solution et il critique la solution proposée afin que les étudiants puissent affiner leurs solutions. L'activité se termine avec la résolution du problème. Les conférences peuvent avoir une structure similaire avec des étudiants présentant un sujet et l'enseignant agissant en tant que « critique » pour les obliger à affiner leurs résultats.

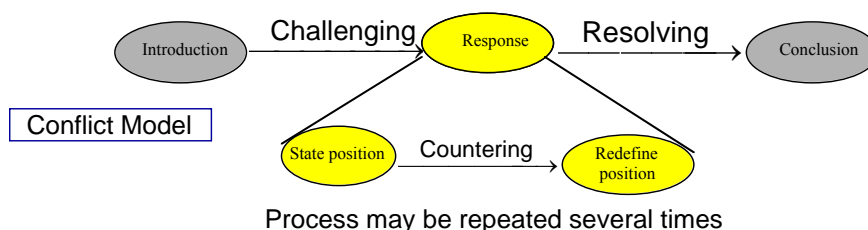


Figure 34 : Structure d'un « conflit »

- Le modèle de la discussion (cf. Figure 35): ce type d'activité est moins structuré avec un certain nombre de scénarios possibles en fonction de qui initialise le processus et qui le continue. Dans ce processus il peut également y avoir des périodes de « conflit » et de désaccord. Le modèle du « conflit » de la section précédente peut être approprié. L'étude dans cette situation est fondée sur la collaboration entre les participants, car ils négocient une compréhension commune du sujet qui constitue le centre de l'activité.

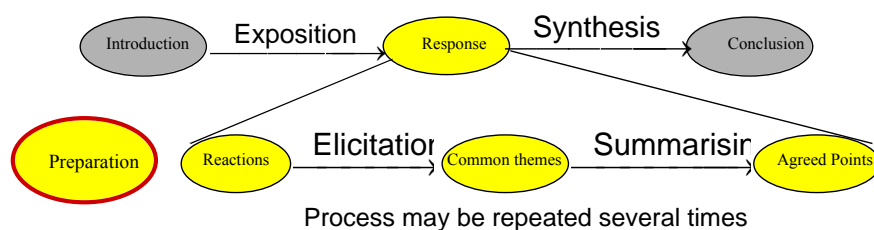


Figure 35: Structure d'un processus de discussion

Le modèle pédagogique s'appuie sur la décomposition des activités pédagogiques en considérant leur but, les types de structure des actions (comme vue précédemment), de l'interaction rhétorique, des outils, des objets et de la situation pratique qui permet à ces intentions de se réaliser. Dans cette approche, les « bonnes » activités pédagogiques sont celles qui sont organisées et mises en application pour permettre aux intentions des enseignants et des étudiants de se réaliser.

Une première étape dans ce processus est de définir les spécifications précises des résultats qui doivent être obtenus par l'activité. En termes pédagogiques, ceci peut s'exprimer principalement en fonction des domaines spécifiques de connaissances associés aux compétences qui devraient être acquises / appropriées / construites. La décomposition des activités en termes de conditions qui sont nécessaires pour réaliser ces intentions implique de décrire une séquence des actions qui doivent être entreprises par les participants pour atteindre les résultats requis.

Le modèle proposé possède six dimensions qui décrivent la décomposition des activités et les conditions associées qui permettent à ces activités de se produire dans une situation pratique. Ces dimensions ont les suivantes

- But de l'activité : les activités sont déterminées par leur but et celles-ci peuvent être exprimées en considérant les objectifs généraux de l'activité, le domaine de connaissance de l'activité et ses résultats.
- Structure de l'activité : des activités peuvent être décomposées en actions requises pour les réaliser et les conditions nécessaires pour que ces actions puissent se produire. Certaines structures d'activités ont été partiellement explicitées ci-dessus.
- Contexte de l'activité : Il est important de préciser les conditions dans lesquelles l'activité se produit : Qui sont les participants ? (Professeurs, étudiants, tuteurs, stagiaires, conférenciers) ; Quels sont les besoins des étudiants en termes de connaissance, compétences et savoir-faire pour participer à l'activité pédagogique ? ; Quel est le niveau de l'activité visée ? (Étudiant préparant une licence, universitaire ou supérieur) ; Où a lieu l'activité ? Quand l'activité va avoir lieu ? ; comment va se dérouler l'activité ? (en ligne, présentiel ou mixte)
- Outils utilisés dans l'activité : ceci peut inclure l'équipement requis pour l'activité ou pour faire des présentations, pour permettre la communication entre les participants à l'activité pédagogique. Elle inclut également les applications de conception et de gestion des cours s'il y en a.

- Objets utilisés dans l'activité : ce sont les ressources utilisées pendant l'activité telle que des textes, des documents, des problèmes et tout autre matériel que les participants vont utiliser pour réaliser l'activité
- Rôles pour les participants de l'activité : Des activités pédagogiques sont structurées par les rôles que les participants adoptent. En particulier, les rôles définissent comment les participants interagissent entre eux pendant les différentes phases de l'activité pédagogique.

Toutes ces dimensions sont décrites par un ensemble de propriétés qui vont être présentes dans l'ontologie de document afin de définir les sous-catégories pertinentes de ressources pour cette application. Mais elles vont aussi être utilisées pour la description de ces ressources dans le schéma de métadonnées. Toutes les propriétés relatives aux dimensions de ce modèle pédagogique sont des extensions de LOM pour la description des ressources. Pour plus de précision sur ce schéma de métadonnées se reporter aux livrables du projet européen CANDLE. Pour la spécification des cours en ligne de CANDLE, le poste de journaliste assisté par ordinateur a été modifié et adapté pour cette nouvelle application.

On peut voir dans les figures suivantes, respectivement une vue globale de l'outil CAT (CANDLE Authoring Tool) (cf. Figure 36), une vue de la partie saisie des métadonnées avec la partie spécifique pédagogie (cf. Figure 38), une vue de la structure de l'activité (cf. Figure 37) et un exemple de page web pour le cours ATM de CANDLE (cf. Figures 39).

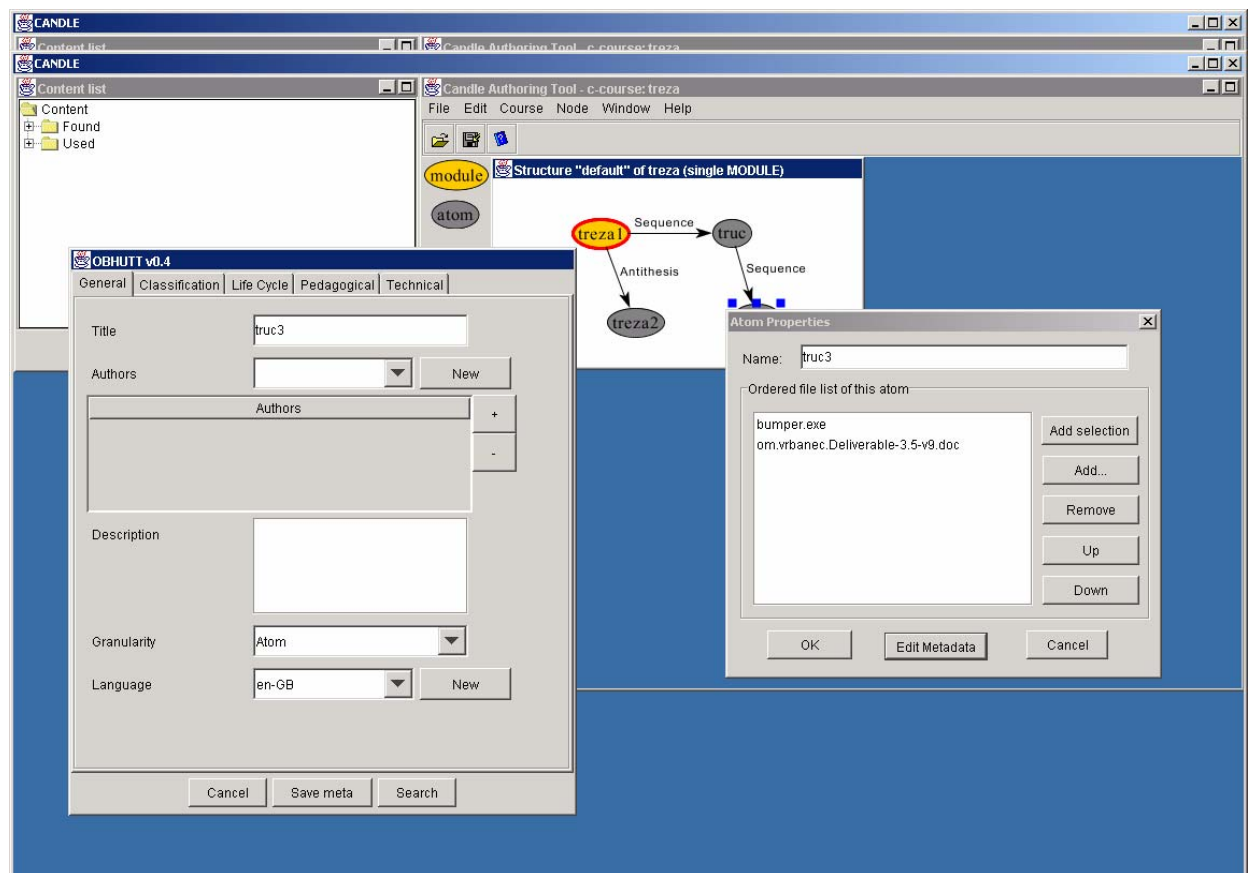


Figure 36 : vue globale de l'outil CAT (CANDLE Authoring Tool)

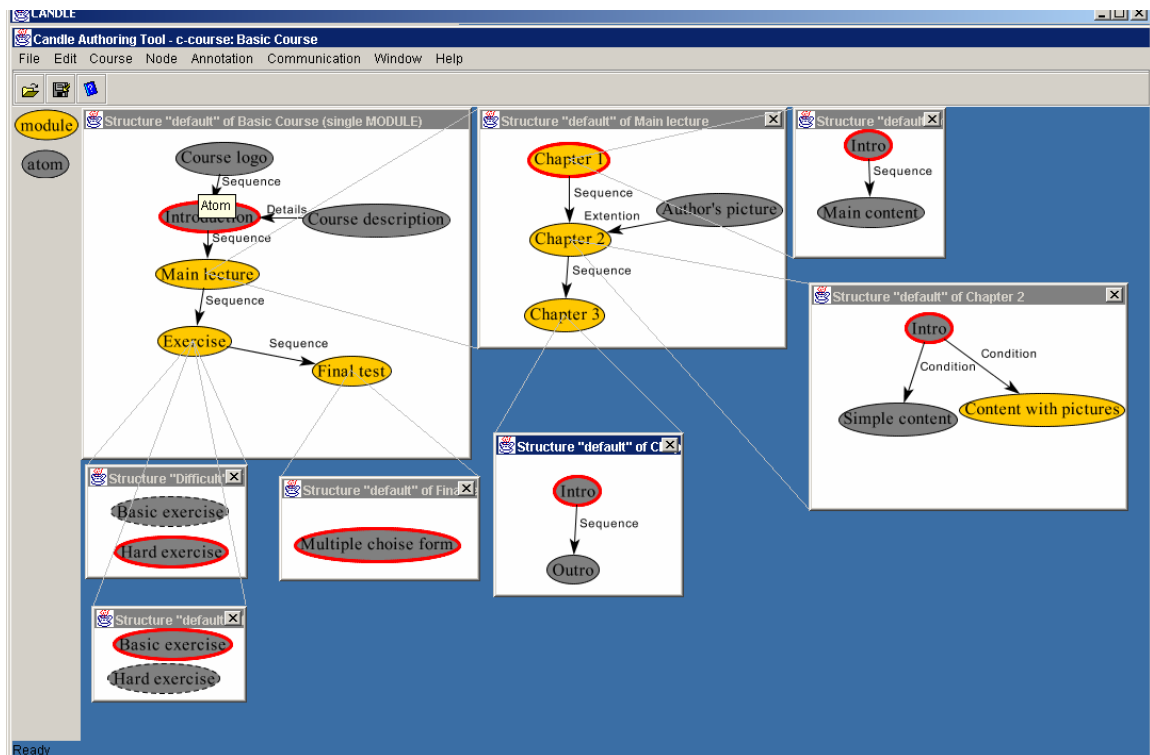


Figure 37 : structure d'une activité avec le CAT

Figure 38 : Saisie des métadonnées dans CAT

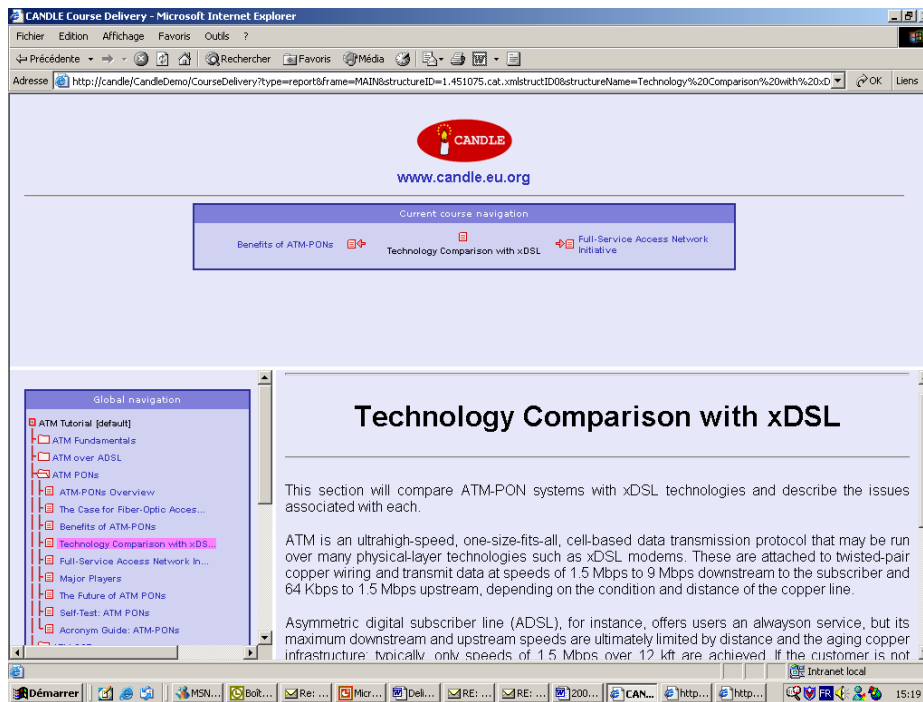


Figure 39 : Le cours CANDLE sur ATM

4.5.3.6 Le Moteur de composition adaptatif

Les quatre mécanismes d'un document virtuel sont mis en œuvre comme suit : la sélection, le filtrage ou adaptation et l'organisation sont traités dans la composition sémantique ; l'assemblage est divisé en une composition logique et une composition de présentation (cf. Figure 40).

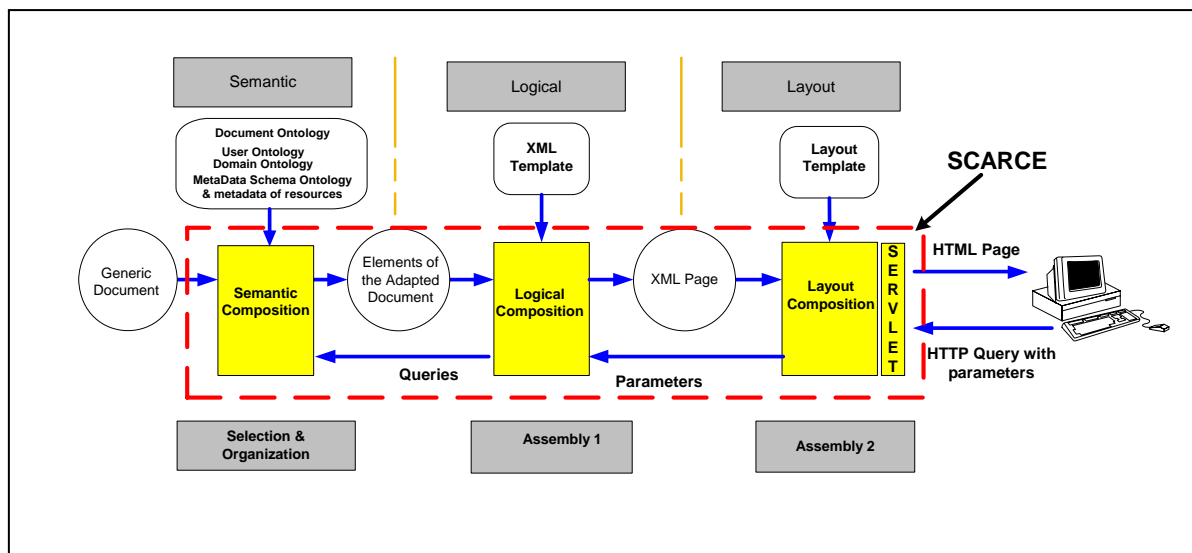


Figure 40 : L'architecture de moteur de composition

Le moteur de composition sémantique génère dynamiquement un graphe sémantique à partir d'un document générique pas à pas – le document réel n'est pas engendré en une seule passe, mais construit au fur et à mesure des interactions avec l'utilisateur. Un document générique se compose d'un espace d'information - un ensemble de fragments, un graphe orienté générique dans lequel les nœuds génériques ont une spécification de contenu pour la sélection et les arcs sont des relations sémantiques, un ensemble de classes d'équivalence et leurs règles d'appartenance, et un stéréotype utilisateur pour chaque technique de navigation adaptative (cf. Figure 41). Une description détaillée du document générique peut être trouvée dans (Garlatti and Iksal 2004). Le graphe sémantique représente une structure narrative adaptée à l'utilisateur courant et liée au contenu correspondant.

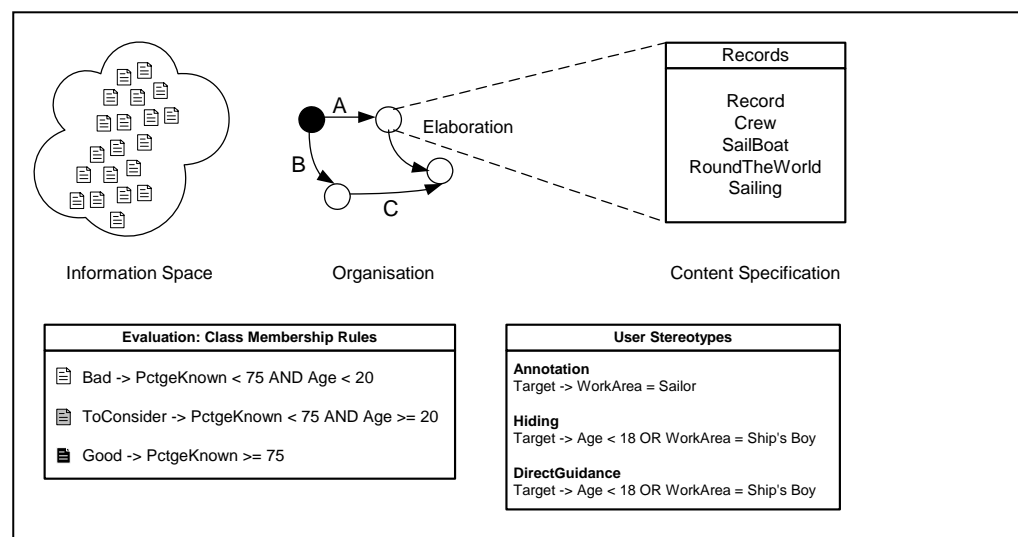


Figure 41 : Un document générique

Le but des moteurs de composition logique et de présentation est de convertir le graphe sémantique en un document hypermédia. La navigation et l'accès aux ressources est fondée sur le graphe sémantique. Le moteur de composition logique envoie des requêtes au moteur de composition sémantique pour construire et parcourir le graphe sémantique en fonction des interactions utilisateur. Pour chaque nœud du graphe sémantique, il calcule une page XML avec ses outils de navigation et le contenu correspondant. Un gabarit XML est utilisé pour obtenir une page XML. Ce gabarit contient notamment toutes les requêtes nécessaires à l'obtention des données dynamiques d'une page XML. Le moteur de composition de présentation produit une page HTML à partir de la page XML en appliquant les règles de présentation fournies dans le gabarit de présentation.

Le moteur de composition sémantique envoie des requêtes au moteur d'inférence de Ontobroker – un système de gestion de connaissances du sémantique web. Il est utilisé pour parcourir et interroger les différentes ontologies et sélectionner les ressources qui s'apparient avec les spécifications de contenu à l'aide des métadonnées incluses dans Ontobroker (Fensel, Angele et al. 1999).

4.5.3.7 La composition adaptative

Une page Web se compose principalement de trois composants dynamiques différents (guides locaux et globaux de navigation et de contenu) (cf. Figure 42). Une étape correspond au calcul de la prochaine page web en fonction de l'interaction utilisateur dans la page web courante. Nous décrivons ici le processus qui permet de passer d'une étape à une autre. Dans la figure 3, le guide de navigation local se compose de trois types différents de liens : pages précédentes, page courante et pages suivantes. Dans cet exemple, l'annotation est utilisée pour la technique de navigation adaptative. Pour calculer la prochaine page XML, Il est nécessaire d'évaluer les ressources correspondant aux pages suivantes si elles n'ont pas encore été calculées. En conséquence, les pages Web suivantes doivent être calculées à l'avance pour pouvoir montrer et annoter ces liens. Le calcul de la page Web suivante est divisé en deux parties : la première est la composition sémantique adaptative et la seconde est la composition logique et de présentation.

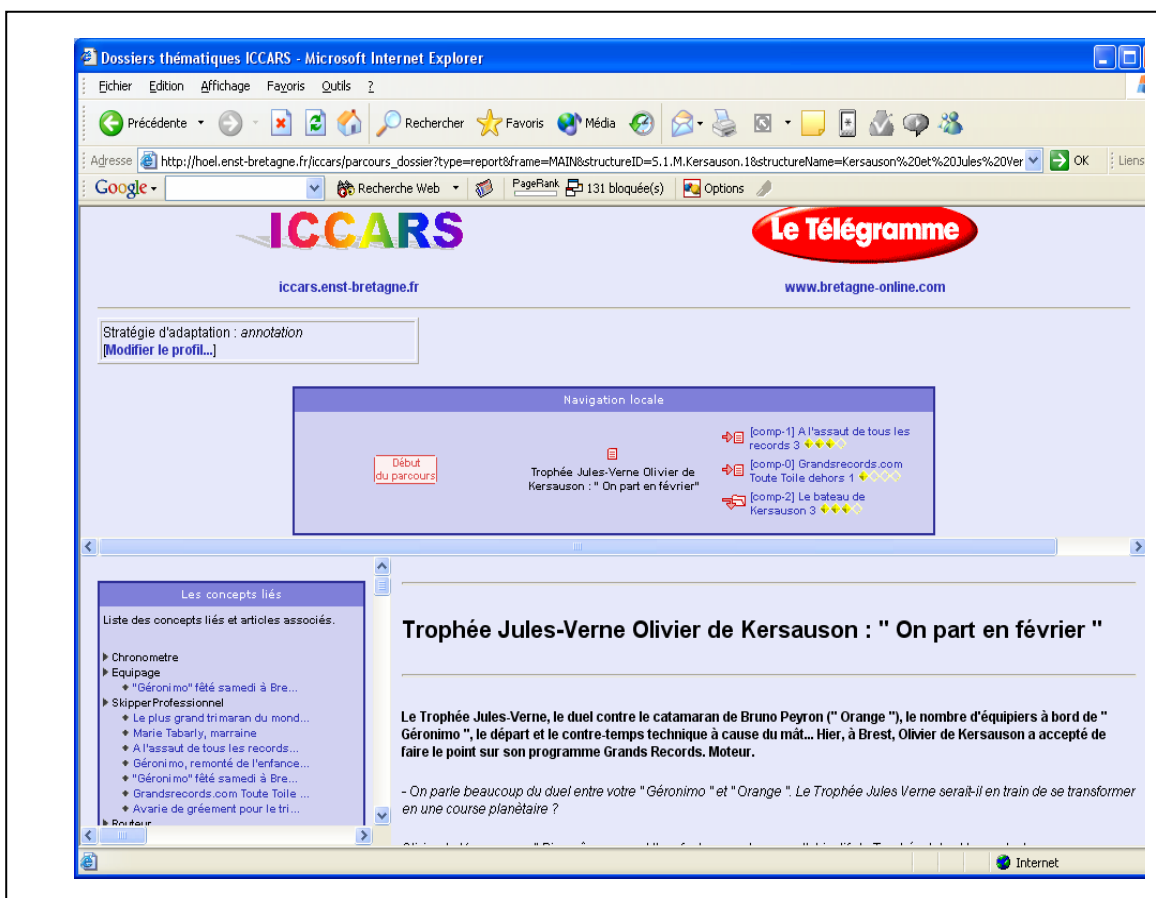


Figure 42 : Disposition de page Web

4.5.3.8 Composition sémantique adaptative

Le moteur de composition sémantique calcule un graphe sémantique pas à pas. Une page web est associée à un nœud du graphe sémantique et généralement, le nœud correspondant à la prochaine page a déjà été calculé pour afficher notamment dans le guide de navigation

locale les liens des pages suivantes. Ce nœud du graphe sémantique est associé à un nœud du graphe générique. A chaque étape, tous les nœuds suivants dans le graphe générique de ce nœud générique sont traités pour rechercher les ressources correspondantes. Pendant l'instanciation et pour chaque nœud suivant, le graphe sémantique est complété de la manière suivante : un nœud générique et sa spécification de contenu pourra engendrer plusieurs nœuds dans le graphe sémantique correspondant, un par ressource sélectionnée, l'arc générique correspondant permet d'engendrer plusieurs arcs - de même nature - dans le graphe sémantique, un par ressource sélectionnée. Pour chaque nœud suivant, le moteur de composition sémantique se compose de trois mécanismes : le premier sélectionne les ressources pertinentes pour le nœud courant, le second évalue ces ressources et les classe dans différentes classes d'équivalence et le troisième détermine les techniques de navigation adaptative autorisées et en applique une.

Dans un premier temps, les spécifications de contenu sont utilisées pour faire une requête au moteur d'inférence – Ontobroker - et les ressources pertinentes sont sélectionnées à partir de l'espace d'information associé au document générique. Les résultats de ce processus sont un ensemble de fragments variants pour le nœud courant. Ces fragments diffèrent sur un sous-ensemble d'index du schéma de métadonnées, par exemple, le niveau de la connaissance pour quelques concepts de domaine, les propriétés techniques, etc. A partir des spécifications de contenu indiquées précédemment (cf. Figure 43), deux fragments variants sont retrouvés. L'un est un fragment atomique et l'autre un fragment abstrait.

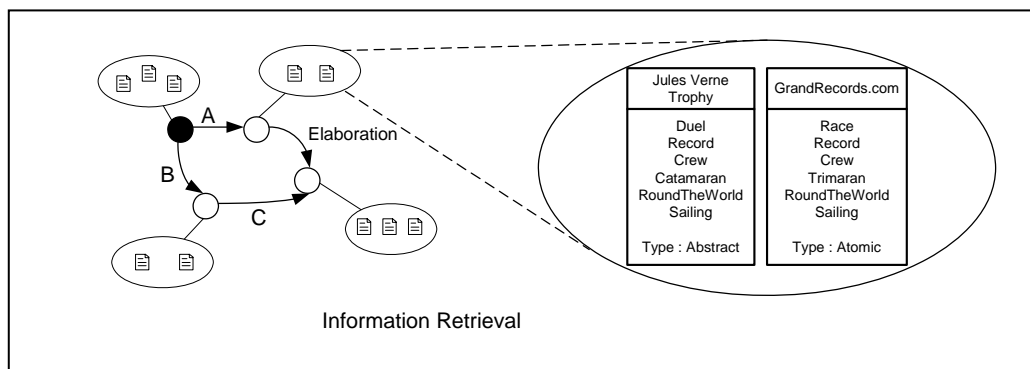


Figure 43 : La sélection

Dans un second temps, tous les fragments variants sont évalués. Pour chaque fragment, le sous-ensemble de métadonnées dédié à l'adaptation et le modèle utilisateur sont utilisés pour décider de l'appartenance aux classes d'équivalence à l'aide des règles d'appartenance. Chaque variante de fragment appartient à une classe et la nouvelle propriété sémantique correspondante est ajoutée au fragment. Cette propriété s'appelle « membre_de » et sa valeur appartient à l'ensemble {« très mauvais », ... , « très bon »}. Cette propriété sera utilisée pour gérer les techniques d'adaptation. Dans la figure 45, nous pouvons voir les résultats de plusieurs évaluations sur différents nœuds, les classifications correspondantes et les résultats de l'application d'une méthode d'adaptation particulière. Les fragments de la figure 44 sont évalués à « bons » et à « mauvais ». L'utilisateur (cf. Figure 44) a 19 ans et il connaît moins de 75% des concepts du fragment pour l'un d'eux et plus de 75% pour l'autre.

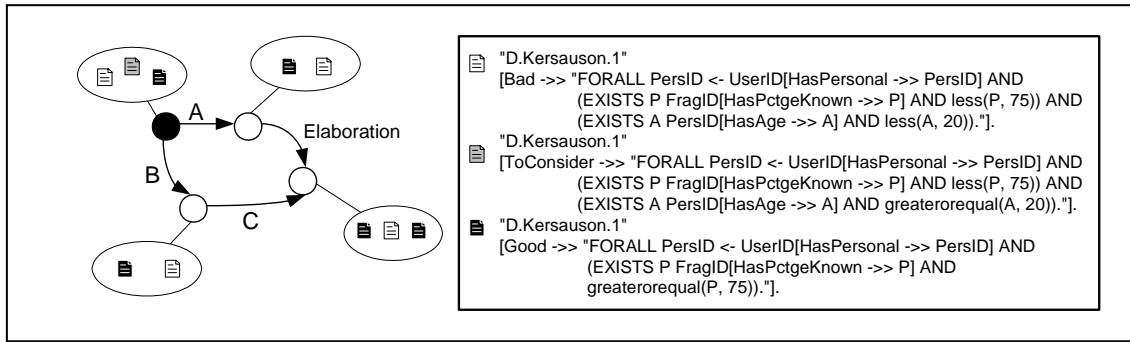


Figure 44 : Évaluation : règles d'appartenance aux classes

Dans un troisième temps, le modèle utilisateur est comparé à tous les stéréotypes associés aux différentes techniques d'adaptation. Ceux qui s'apparient au modèle d'utilisateur sont autorisés pour l'utilisateur courant. Dans la figure 45, seuls les stéréotypes des techniques de masquage et de guidage direct s'apparient au modèle utilisateur. A partir des préférences de l'utilisateur, la première technique dans l'ordre total, donné par les préférences de l'utilisateur, est choisie. La figure 46 nous montre l'application du masquage.

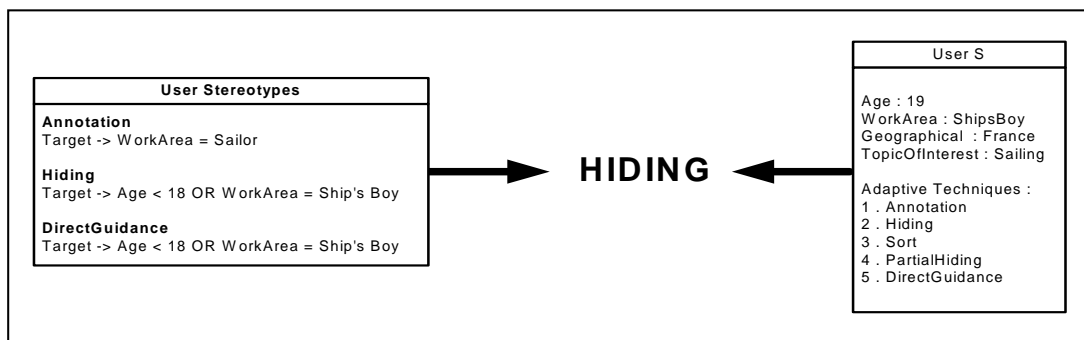


Figure 45 : Choix d'une méthode d'adaptation

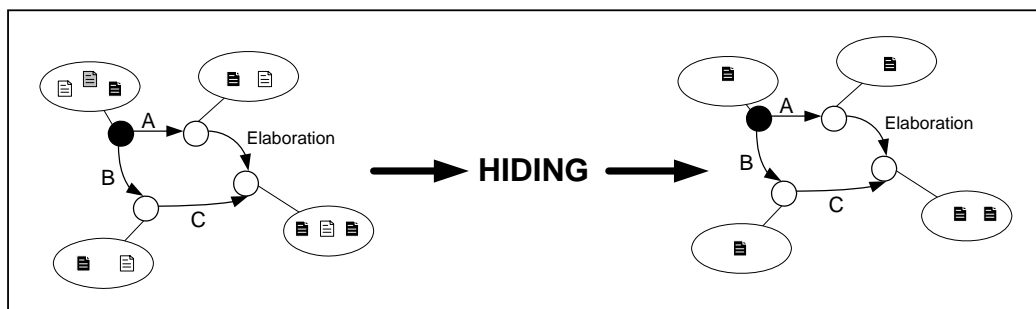


Figure 46 : Application d'une technique adaptative

En conclusion, un fragment variant est maintenue ou pas dans le nœud courant en fonction de la technique d'adaptation utilisée. Par exemple, si on permet à un utilisateur d'avoir les techniques d'annotation, de masquage et de guidage direct, tous les fragments variants sont

conservés quelque soit leur classe d'équivalence. En effet, il est nécessaire de garder tous les fragments pour pouvoir les annoter. Au contraire, si on autorise seulement à un utilisateur le masquage et le guidage direct, uniquement les fragments variants les plus pertinents sont conservés (cf. Figure 46). Les autres fragments variants sont supprimés puisque l'utilisateur ne peut pas y accéder. Quelques techniques de navigation adaptative, comme le masquage, le masquage partiel ou le guidage direct permettent la suppression des fragments variants non pertinents si ce sont les seules techniques autorisées. Elles ont donc des conséquences directes sur le contenu et la structure de document et sur l'adaptation de contenu. Si quelques fragments variants sont supprimés, le contenu et la structure de document peuvent être modifiés.

4.5.3.9 Composition logique et composition de présentation

Selon l'interaction avec utilisateur, le moteur de servlet (cf. Figure 40) reçoit la requête HTTP avec des paramètres et les transmet au moteur de composition de présentation. Le moteur de composition de présentation envoie ces paramètres au moteur de composition logique. Par exemple, nous supposons que l'utilisateur demande les prochaines pages web au moyen du guide local de navigation. Le lien muni de ces paramètres est explicité dans la figure 47.

```
« http://hoel.enst-bretagne.fr/iccars/parcours_dossier?type=report&frame=
MAIN&structureID=S.1.M. Kersauson.1&structureName=Kersauson et Jules
Verne&componentID=C.3.S.1.M. Kersauson.1-Iksal&fragmentID=
F. Kersauson.3-Iksal »
```

Figure 47 : paramètre reçu par le moteur de servlet

Cette requête HTTP se compose de six paramètres :

1. L'URL du document virtuel courant :
2. « http://hoel.enst-bretagne.fr/iccars/parcours_dossier ? »
3. Le gabarit XML pour la prochaine page Web : « type = report »
4. La partie de page web qui sera changée : « frame = MAIN »
5. Données (identification et nom) de la structure narrative choisie par l'utilisateur :
6. « structureID=S.1.M. Kersauson.1&structureName=Kersauson et Jules Verne »
7. Le prochain noeud : « componentID=C.3.S.1.M. Kersauson.1-Iksal »
8. Le fragment correspondant : « fragmentID=F. Kersauson.3-Iksal »

Au moyen de ces paramètres, le moteur de composition logique peut choisir le gabarit XML pour le document courant (paramètres 1 et 2). Le gabarit XML contient la structure logique et un ensemble de requêtes pour générer une nouvelle page XML. Ce moteur construit

cette page XML qui se compose du contenu, d'une description des guides de navigation et également de quelques données relatives à la technique d'adaptation et l'état de l'évaluation de chaque lien. Les paramètres 4 et 5 sont utilisés pour obtenir le prochain nœud dans le graphe sémantique et tous ses voisins directs dans le graphe générique correspondant. Le paramètre 6 est utilisé pour obtenir le contenu de la page Web. Les paramètres 4, 5 et 6 sont utilisés pour faire des requêtes au moteur de composition sémantique pour instancier le gabarit XML. A partir de la page XML et d'un gabarit de présentation, le moteur de composition de présentation peut produire une page HTML.

Dans des guides de navigation, les liens ont des propriétés pour la gestion de l'adaptation (Brusilovsky 1996). La figure 48 montre le guide de navigation locale dans la page XML avec des propriétés de lien pour l'adaptation. Ce guide se compose de liens vers les pages précédentes, la page courante et les pages suivantes. Pour chaque lien, il y a une propriété donnant l'état d'évaluation donné par le moteur de composition sémantique qui sera traité par le moteur de composition de présentation pour afficher chaque état particulier sur la page Web (cf. Figure 48).

```

locale</component_title> de <component_title>Navigation
structureID= de Jules Verne de structureName= de " S.1.M. Kersauson.1
" strategy= " annotation " de <local_navigation de " Kersauson et " >
<previous>
<begin/>
</previous>
>Trophée <current Jules-Verne Olivier de Kersauson de " CM_Molecule "
de type= : ' ' Sur < de ' ' de févrie-r d'en de partie/current>
<next>
componentID= fragmentID= " F. Kersauson.3-Iksal " de
<next_article_link de " C.3.S.1.M. Kersauson.1-Iksal " >
(Nanard and J. Nanard; Nanard and J. Nanard) Disques tous de les du
l'assaut un De
<evaluation_state>3</evaluation_state>
</next_article_link>
componentID= fragmentID= " F. Kersauson.7-Iksal " de
<next_article_link de " C.2.S.1.M. Kersauson.1-Iksal " >
[comp-0] Dehors de Grandsrecords.com Toute Toile
<evaluation_state>1</evaluation_state>
</next_article_link>
go_structureID= from_componentID= from_fragmentID= " F.M.
Kersauson.2-Iksal " de <next_report_link de " S.1.M. Kersauson.2 " de
" C.4.S.1.M. Kersauson.1-Iksal " >
[comp-2] Le bateau de Kersauson
<evaluation_state>3</evaluation_state>
</next_report_link>
</next>
</local_navigation>
</component>
</panel>
</page>

```

Figure 48 : Guide local de navigation : composant XML

4.5.4 Conclusion

Dans ces nouveaux systèmes d'information, les informations ne sont plus uniquement disponibles localement mais sont distribuées sur Internet. Leur partage et leur réutilisation deviennent nécessaires pour la conception de ces systèmes d'information. Il n'est pas toujours facile d'assurer leur réutilisation par des machines et ceci d'autant plus qu'il y a un problème de complexité liée à la masse d'information disponibles et à leur hétérogénéité puisqu'ils sont créés par différents auteurs et peuvent être décrites – indexées - par différentes personnes. Cette réutilisation implique une phase de recherche d'information et une génération dynamique des systèmes d'information. Il y a donc des problèmes de recherche d'information, d'automatisation ou de semi automatisation de certaines tâches et d'interopérabilité à résoudre. Dans notre contexte, les applications ont au moins deux catégories d'utilisateurs : des lecteurs qui ont pour tâches principales la lecture et la compréhension d'un document hypermédia adapté à leurs besoins ; des auteurs qui possèdent des compétences et des savoir-faire leur permettant de spécifier les propriétés nécessaires à la génération de documents adaptés à ces lecteurs. Les auteurs appartiennent à une communauté de pratiques : ils partagent des connaissances communes qui leur permettent de travailler ensemble. Ce sont les connaissances explicites de ces communautés de pratique qui sont la clé pour la conception et l'automatisation de la génération de ces systèmes d'information.

Nous avons proposé un moteur de composition sémantique appelé SCARCE – SemantiC & Adaptive Retrieval and Composition Engine) et une méthodologie de conception de systèmes d'information sémantiques et adaptatifs dont les propriétés principales sont les suivantes :

- Les connaissances explicites des communautés de pratiques sont représentées par des ontologies, des schémas de métadonnées sémantiques et de la logique. Elles servent de fondement aux modèles proposés. Ces connaissances permettent aussi d'assurer la cohérence des résultats et de fournir une méthodologie de conception de tels systèmes.
- Les mécanismes de sélection, d'organisation et d'adaptation peuvent être spécifiés au niveau sémantique par des paramètres déclaratifs issus des connaissances explicites de la communauté de pratiques pour assurer une flexibilité maximale du moteur de composition.
- Les ontologies sont organisées en quatre catégories l'une regroupant le modèle de document et le modèle d'adaptation, une ontologie du domaine, une ontologie des utilisateurs et une autre pour le schéma de métadonnées. Elles sont faiblement couplées et donc aisément modifiables. Une grande partie de celles-ci sont des paramètres du moteur de composition et peuvent donc changer et/ou être mis à jour pour la maintenance et/ou la création de nouvelles applications.

Ce moteur de composition possède deux niveaux de paramétrages qui assurent une plus grande flexibilité et extensibilité : d'une part les quatre ontologies et d'autre part le document générique - qui est une spécification. La maintenance et l'évolution d'un système d'information sont assurées par la modification d'un document générique dont on peut changer certaines spécifications : sélection, organisation ou adaptation.

La création d'un nouveau système d'information peut se faire par modification et/ou changement de l'une des ontologies ou des quatre. Il est bien évident que ceci n'est possible que si les principes fondamentaux de ce moteur de composition sont pertinents pour le nouveau système d'information. Ces principes fondamentaux sont d'une part ceux de l'ontologie de document - partie non dépendante de l'application et d'autre part ceux des règles de composition utilisées à l'exécution -, c'est-à-dire celles qui permettent de transformer le document générique en un graphe sémantique. Ces deux aspects sont dépendant l'un de l'autre car l'ontologie de document est porteuse de sens qui doit être respecté par les règles de composition. Par exemple, pour ICCARS et CANDLE, le document générique sous-tend une tâche principale de lecture et le graphe sémantique engendré n'impose pas d'ordre entre les nœuds suivants le nœud courant et le nombre de ressources accessibles à partir d'un nœud générique est volontairement limité. En effet, pour augmenter l'uniformité du document fourni, nous avons choisi d'associer un espace d'information à un document. L'idée principale est de s'assurer que nous avons le bon contenu au bon endroit pour l'utilisateur droit. Il sera intéressant de libérer cette contrainte selon les buts d'utilisateur et les services que nous voulons proposer. Dans le cas d'un système de recherche documentaire typique, il est nécessaire de fournir le maximum du fragment approprié, mais dans le cas d'une application pédagogique, il est nécessaire d'assurer la cohérence entre tous les fragments recherchés. Nous n'avons pas de solution à l'heure actuelle ; c'est un problème intéressant qui doit être étudié à l'avenir.

L'environnement SCARCE a aussi été utilisé dans un projet appelé KMP pour « Knowledge Management Portal ». Ce projet est le fruit d'une réflexion menée dans le cadre du Laboratoire des Usages de Sophia Antipolis, dont la spécificité est de se focaliser sur l'observation des usages TIC dans le domaine de l'E-Business. Au regard des travaux existants, la contribution de ce projet est triple : i) Rendre compte des pratiques de « Knowledge Management » à l'échelle d'un réseau inter-firmes et inter-institutions ; ii) Analyser les possibilités et opportunités de formalisation des pratiques d'échange et de création de connaissances inter-firmes, qui sont essentiellement informelles ; iii) Concevoir, à partir de l'analyse et de l'anticipation des usages, le média le mieux approprié à ces pratiques. Dans le cadre du projet KMP, l'utilisateur doit d'une part réaliser de nombreuses tâches différentes et d'autre part faire de la recherche d'information interactive. Il n'était donc plus possible de conserver le modèle de document actuel parce qu'il est nécessaire de modéliser les tâches utilisateur par un modèle de tâches hiérarchiques comme dans le projet SWAN et la requête d'un nœud générique peut sélectionner une centaine de ressources. La stratégie de génération du graphe sémantique n'est plus applicable. Il n'est en effet pas pensable de générer autant de nœud dans le graphe qu'il a de ressources sélectionnées. Il est donc nécessaire de changer le modèle de document mais aussi le modèle d'adaptation. Le sens d'un modèle de tâches hiérarchiques est très différent de celui du modèle de document d'ICCARS et de CANDLE. Les règles du moteur de composition doivent donc aussi changer. Pour KMP, il est par contre possible d'intégrer les deux modèles de document en faisant l'hypothèse que celui d'ICCARS devient une tâche particulière qui est une tâche de lecture en ayant bien évidemment un moteur de composition possédant les deux jeux de règles l'un pour le modèle de tâches hiérarchiques et l'autre pour l'autre le modèle de document d'ICCARS.

4.6 Crédits

Les études menées dans le cadre du projet ICCARS et CANDLE doivent beaucoup à l'équipe du Télégramme participant au projet dont notamment M. Jean-Yves Chalm, M. Olivier Clech et René Le Clech, à la société Atlantide et plus particulièrement à M. Patrick Poupon, à M. Philippe Kervella et M. Didier Riou, aux différents stagiaires et CDD qui ont travaillé sur ce projet, à M. Philippe Tanguy qui a développé les moteurs de composition logique et de présentation, à M. Sébastien Iksal qui a effectué sa thèse au travers de ces projets et qui a eut la lourde tâche d'acquérir les connaissances communes des journalistes et de réaliser le moteur de composition sémantique, sans oublier toutes les discussions et échanges que nous avons eu tout au long de sa thèse et plus tard. Ces études ont aussi bénéficié de tout le travail réalisé en commun avec M. Yvon Kermarrec et toute l'équipe qui nous entourait pour les deux projets. Pour CANDLE, ce travail n'aurait jamais pu se faire sans tous les membres du consortium de ce projet européen.

Ces études ont pu être effectuées grâce aux discussions et échanges réalisés dans le groupe « Document Virtuel » français et plus particulièrement M. Marc Nanard, Mme Jocelyne Nanard, M. Michel Crampes et Mme Sylvie Ranwez.

Ces études ont bénéficié de l'apport du travail réalisé dans le cadre de l'action spécifique Web Sémantique du RTP 33 « Documents et contenu : création, indexation, navigation » et plus particulièrement de M. Jean Charlet, Mme Chantal Reynaud et M. Yannick Prié et aussi de l'apport du RTP 39 « Apprentissage, éducation et formation » et plus particulièrement de M. Nicolas Balacheff, M. Alain Derycke, M. Pierre Tchounikine et toute l'équipe du comité de pilotage.

4.7 Bibliographie

- Aroyo, L., P. De Bra, et al. (2003). Embedding Information Retrieval in Adaptive Hypermedia: IR meets AHA!. AH 2003: Workshop on Adaptive Hypermedia And Adaptive Web-Based Systems, Technische Universiteit Eindhoven.
- Baget, J.-F., E. Canaud, et al. (2005). "Les langages du Web sémantique." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.
- Balpe, J. (1997). Trois mythologies et un poète aveugle. Colloque "De la différence des arts", Paris, IRCAM.
- Berners-Lee, T. (1998). Why RDF model is different from the XML model. <http://www.w3.org/DesignIssues/RDF-XML.html>, W3C.
- Berners-Lee, T., J. Hendler, et al. (2001). "The semantic web." The Scientific American 5(284): 34-43.
- Bra, P. D. (1999). Design issues in adaptive hypermedia application development. Workshop on Adaptive Systems and User modeling on the World Wide Web, Toronto and Banff.
- Brézillon, P. (2002). Hors du contexte, point de salut. Séminaire "Objets Communicants", Autrans, France.
- Brusilovsky, P. (1996). "Methods and techniques of adaptive hypermedia." User Modeling and User-Adapted Interaction 6(2-3): 87-129.

- Brusilovsky, P., E. Schwarz, et al. (1996). ELM-ART: An intelligent tutoring system on World Wide Web. Third International Conference on Intelligent Tutoring Systems, ITS-96, Montreal.
- Brusilovsky, P. and H. Su (2002). Adaptive Visualization Component of a distributed Web-base Adaptive Educational System. ITS 2002, Biarritz, Springer Verlag.
- Calfee, R. and M. J. Chambliss (1987). "The structural design features of large texts." Educational Psychologist 3-4(22): 357-378.
- Charlet, J. (2002). L'Ingénierie des connaissances : développements, résultats et perspectives pour la gestion des connaissances médicales. Informatique. Paris, Université Pierre et Marie Curie.
- Charlet, J., B. Bachimont, et al. (2005). "Ontologies pour le web sémantique." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.
- Charlet, J., P. Laublet, et al., Eds. (2003). Action spécifique 32 CNRS/STIC, Web Sémantique, rapport final.
- Charlet, J., P. Laublet, et al. (2005). "Le Web sémantique." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.
- Chevrin, V., J. Rouillard, et al. (2006). Multi-channel Adaptive contexte-sensitive intermediation : Reconciling Marketing goals and User Needs. Workshop MAC, Glasgow, Scotland, UK.
- Christophides, V. (1998). Electronic Document Management Systems. <http://www.ics.forth.gr/~christop/>, UCH/FORTH.
- Coirier, P., D. Gaonac'h, et al. (1996). Psycholinguistique textuelle : Approche cognitive de la compréhension et de la production des textes, Puf.
- Conlan, O., D. Lewis, et al. (2003). Applying Adaptive Hypermedia Techniques to Semantic Web Service Composition. Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems in The Twelfth International World Wide Web Conference, Budapest, Hungary.
- Crampes, M., S. Ranwez, et al. (2000). Ontology-supported and ontology-driven conceptual navigation on the world wide web. Proceedings of Hypertext 2000, San Antonio, USA, ACM Press.
- Denhière, G. and S. Baudet (1992). Lecture, compréhension de texte et science cognitive. Paris, PUF.
- Dijk, T. A. V. and W. Kintsch (1983). Strategies of discourse comprehension. Orlando, Academic Press.
- Ehrlich, M.-F., A. Charles, et al. (1992). La superstructure dans les textes expositifs est-elle prise en charge lors de la sélection des informations importantes ? Le résumé de texte. M. Charolles and A. Petitjean: 183-202.
- Englert, C. S., S. R. Stewart, et al. (1988). "Young writers' use of text structure in expository text generation." Journal of Educational Psychology(80): 143-151.
- Falquet, G., L. Nerima, et al. (1999). Des documents virtuels pour lire les bases de données. Workshop sur les Documents Virtuels Personnaliables, IHM'99,, Montpellier.
- Fayol, M. (1985). Le récit et sa construction. Neuchatel, Paris, Delachaux & Niestlé.
- Fayol, M. and J.-M. Monteil (1988). "The notion of "script" : from general to developmental and social psychology." Cahiers de Psychologie Cognitive(8): 335-361.

- Fensel, D., J. Angele, et al. (1999). On2broker: Semantic-Based Access to Information Sources at the WWW. World Conference on the WWW and Internet, WebNet 99, Honolulu, Hawaii, USA.
- Folman, S. and S. Gissi (1990). "Intercultural rhetorical differences in meaning construction." Communication & Cognition 23(1): 45-92.
- Garlatti, S. and S. Iksal (1999). Documents virtuels personnalisables pour des systèmes d'informations en ligne. Workshop sur les Documents Virtuels Personnalises, IHM'99,, Montpellier.
- Garlatti, S. and S. Iksal (2000). Concept Filtering and Spatial Filtering in an Adaptive Information System. Adaptive Hypermedia and Adaptive Web-Based Systems. P. Brusilovsky, O. Stock and C. Strapparava. Trento, Springer Verlag: 315-318.
- Garlatti, S. and S. Iksal (2004). A Flexible Composition Engine for Adaptive Web Sites. Adaptive Hypermedia and Adaptive Web-Based Systems, AH 2004. P. D. Bra and W. Nejdl, Springer Verlag. LNCS 3137: 115-125.
- Garlatti, S., S. Iksal, et al. (1999). Adaptive On-line Information System by means of a Task Model and Spatial Views. Second Workshop on Adaptive Systems and user Modeling on the World Wide Web, Toronto and Banff, Eindhoven University of Technology.
- Garlatti, S. and Y. Prié (2005). "Adaptation et personnalisation dans le sémantique web." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.
- Geurts, J. P., S. Bocconi, et al. (2003). Towards ontology-driven discourse: from semantic graphs to multimedia presentations, Centrum voor Wiskunde en Informatica.
- Gruber, T. R. (1993). Toward principles for the design of ontologies used for knowledge sharing. International Workshop on Formal Ontology, Kluwer Academic.
- Gruber, T. R., S. Vemuri, et al. (1996). Model-based Virtual Document Generation, Knowledge Systems Laboratory.
- Iksal, S. (2002). Spécification Déclarative pour des documents virtuels personnalisables. IASC. Brest, ENST Bretagne.
- Iksal, S. and S. Garlatti (2001). Revisiting and Versioning in Virtual Special Reports. Hypermedia: Openness, Structural Awareness and Adaptivity, Third Workshop on Adaptive Hypertext and Hypermedia. S. Reich, M. M. Tzagarakis and P. D. Bra. Århus, Denmark, Springer Verlag: 264-279.
- Isaac, A., P. Courounet, et al. (2004). Un système d'annotation multiforme et communautaire de documents audiovisuels : OPALES. Journée sur les Modèles Documentaires de l'Audiovisuel, Semaine du Document Numérique, La Rochelle.
- LTSC (1999). Learning Object Metadata Working draft document 3. <http://ltsc.ieee.org/doc/wg12/LOM-WD3.htm>, Learning Technology Standard Committee.
- Maedche, A., S. Staab, et al. (2001). Seal - a framework for developping semantic portals. BNCOD, British National Conference on Databases, Oxford, UK, Springer Verlag.
- Mann, W. C. and S. A. Thompson (1988). "Rhetorical Structure Theory: Toward a functional theory of text organization." Text 8(3): 243-281.
- Mathews, S. R. (1982). The impact of prior knowledge on accessibility and availability of information from prose. Discourse Processing. A. Flammer and W. Kintsch. Amsterdam, North-Holland.
- Milosavljevic, M., F. Vitali, et al. (1999). Introduction of the virtual document workshop. Workshop on Virtual Document, Hypertext Functionality and the Web, Toronto.

- Minski, M. (1975). A framework for representing knowledge. The psychology of computer vision. P. Winston. New York, Mc Graw Hill.
- Nanard, J. and M. Nanard (1995). "Hypertext Design Environments and the Hypertext Design Process." Communication of the ACM 38(8): 49-56.
- Nanard, M. and J. Nanard (1989). Macweb: un outil pour élaborer des documents. Workshop on Object Oriented Document Manipulation, WOOD-MAN'89.
- Nanard, M. and J. Nanard (2004). Vers une spécification par intention de la production de documents. Intelligence collective : partage et redistribution des savoirs,, Nîmes.
- Nanard, M. and J. Nanard (2005). Saphir : un cadre de référence pour une spécification des hypermédias par intention. H2PTM, Hypermédias Hypertexts Products Tools and Methods, Lavoisier.
- Oberlander, J. and M. Milasavljevic (1997). Flexible Hypertext - introduction. Hypertext'97, Proceedings of Flexible Hypertext Workshop.
- Paris, C., M. F. Wu, et al. (2004). Myriad: An Architecture for Contextualized Information Retrieval and Delivery. Proceedings of the 3th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems. Eindhoven, The Netherlands, Springer: 205-214.
- Perfetti, C. A., J.-F. Rouet, et al. (1999). Toward a theory of documents representation. The construction of mental representations during reading. H. v. Oostendorp and S. R. Goldman. Mahwah, NJ, Lawrence Erlbaum Associates: 99-122.
- Prié, Y. and S. Garlatti (2005). "Métadonnées et annotations dans le sémantique web." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.
- Ranwez, S. (2000). Composition Automatique de Documents Hypermédia Adaptatifs à partir d'Ontologies et de Requêtes Intentionnelles de l'Utilisateur. Montpellier, Université de Montpellier 2.
- Ranwez, S. and M. Crampes (1999). Conceptual document and hypertext documents are two different forms of virtual document. Workshop on Virtual Document, Hypertext Functionality and the Web, Toronto.
- Rekik, Y. A., C. Vanoirbeek, et al. (1999). Reusing Dynamic Document Fragment through Virtual Documents: Key Issues In Document Engineering. Workshop sur les Documents Virtuels Personnalises, IHM'99, Montpellier.
- Rich, E. (1989). Stereotypes and user modeling. user models in dialog systems. A. Kobsa and W. Wahlster. Berlin, Springer Verlag: 35-51.
- Rouet, F.-F., M. Favard, et al. (1997). "Studying and using multiple documents in History: Effects of discipline expertise." Cognition and Instruction 15(1): 85-106.
- Schank, R. C. and R. P. Abelson (1997). Scripts, plans, goals and understanding. Hillsdale, NJ : Erlbaum.
- Stevenson, I. (2003). Pedagogical Model for Candle. Technical Report deliverable WP6 of IST CANDLE Project.
- Tazi, S. and Y. Altawki (1999). Création de documents virtuels - Cas des supports de cours. Workshop sur les Documents Virtuels Personnalises, IHM'99,, Montpellier.
- Vercoustre, A.-M., J. Dell'Oro, et al. (1997). Reuse of Information through virtual documents. Second Australian Document Computing Symposium, Melbourne, Australia.

- Watters, C. and M. Shepherd (1999). Research issues for virtual documents. Workshop on Virtual Document, Hypertext Functionality and the Web, Toronto.
- Wenger, E. (1998). Communities of Practice - Learning, Meaning and Identity. Cambridge, Cambridge University Press.

Chapitre 5. Environnement informatique d'apprentissage humain

5.1 Préambule

Cette quatrième étape est la suite de l'étape précédente puisqu'elle en possède presque toutes les caractéristiques. En effet elle a pour objet la composition sémantique et adaptative de cours en ligne fondée sur des pratiques et des théories didactiques. Mais par rapport à l'étape précédente, l'acquisition des connaissances explicites de la communauté de pratiques est plus difficile. En effet, l'introduction de nouveaux environnements informatiques d'apprentissage humain ou EIAH va profondément modifier les pratiques courantes. Il va donc être nécessaire de mettre en œuvre une méthodologie d'acquisition de ces connaissances adaptée à ce contexte. Cette dernière doit aussi s'intéresser à l'évaluation des résultats pour valider et/ou modifier les pratiques engendrées et donc ces connaissances explicites. Le principal projet de cette étape s'appelle MODALES pour MOdelling Didactic-based Active Learning Environment in Sciences. Ce projet a pour objet la mise en place de séquences de formation, à destination de futurs enseignants (PLC2 et PE), (appelés ci-après apprenants) fondées sur l'utilisation de ressources multimédias et s'appuyant sur des savoir-faire d'experts en éducation et sur la modélisation de pratiques réelles. Il s'agit de mettre en place des scénarios didactiques qui permettent à de futurs enseignants « d'apprendre leur métier » en réalisant des séquences d'enseignement à partir d'une approche de type pédagogie active. La

conception d'un EIAH nécessite de mettre en œuvre des théories didactiques, des théories de l'activité humaine, ainsi que des modèles et théories informatiques. Se pose alors la question de l'articulation entre la modélisation des situations et leur transposition informatique. Ce projet est en cours de réalisation et se compose de trois étapes dont une seule a été achevée actuellement. Ce chapitre présente brièvement quelques résultats de la première étape. Nous avons mis en œuvre une méthodologie de co-conception de scénarii didactiques fondée sur des pratiques réelles et une approche transdisciplinaire comme fondement d'un EIAH.

La plate-forme développée des projets CANDLE et ICCARS sert de point de départ à ce projet et sera étendue. L'objectif est de mettre l'apprenant dans un contexte d'apprentissage actif (pédagogie active ou apprentissage par problèmes ou problem-based learning) lors de la résolution de problèmes dans le cadre d'une formation en ligne. Il mettra l'apprenant en situation d'éprouver sa capacité à analyser un problème, à bâtir une stratégie de résolution, à opérer des choix par rapport au problème à résoudre : choix de la problématique, sélection des sources documentaires, modalité de sélection de ces sources. L'étude du savoir-faire des enseignants en termes de pédagogie active et de didactique des sciences servira de fondements pour la structuration des cours personnalisés. Ces principes de structuration et une sélection des ressources à disposition des enseignants permettront la génération dynamique des cours en utilisant les principes des documents virtuels adaptatifs. La structuration des cours et les contenus associés seront étudiés pour différentes catégories d'apprenants afin d'analyser les variations interindividuelles et inter catégories en terme d'apprentissage. Il s'agira ici d'obtenir les paramètres pertinents qui permettront d'adapter les cours aux différentes catégories d'apprenants et dans une catégorie en fonction des individus. Il est bien évident que ces critères seront validés par une évaluation de la qualité d'apprentissage individuelle et par catégorie.

Ce paragraphe est principalement extrait du rapport annuel du projet MODALES réalisé par les différents partenaires et lus particulièrement par Sylvain Laubé de l'IUFM. Ce projet étant actuellement en cours de réalisation, ce rapport est une excellente synthèse des travaux réalisés sur ce projet, même si la partie dédiée à la psychologie cognitive n'y est pas intégrée. Il s'organise de la manière suivante : dans un premier temps, le contexte du projet MODALES est présenté. Dans un second temps, les problématiques et la méthodologie employée sont explicitées. Puis, les différentes étapes d'élaboration des scénarios par les formateurs sont détaillées. Ensuite, les scénarios sont formalisés par des modèles de tâches hiérarchiques. Ce paragraphe se termine par une conclusion et une bibliographie.

5.2 Introduction

La création des CAREST (Centres d'Autoformation et de Ressources en Sciences et Technologie) au sein de l'IUFM de Bretagne s'est traduit dans les plans de formations 2004-2007 par une demande institutionnelle de nouveaux dispositifs didactiques à destination des maîtres en formation initiale et continue, dispositifs reposant sur l'autoformation par l'usage des TIC et visant à permettre l'évolution des connaissances en sciences et en didactique des sciences. Les travaux de recherche liés à la conception et l'utilisation de tels Environnements Informatiques pour l'Apprentissage Humain (EIAH) rentre dans le cadre de la recherche en didactique des sciences et de l'« Ingénierie des EIAH », car il s'agit bien de définir des concepts, méthodes et techniques reproductibles et/ou réutilisables facilitant la mise en place (conception – réalisation – expérimentation – évaluation - diffusion) d'environnements de

formation ou d'apprentissage en permettant de dépasser le traitement ad hoc des problèmes (Tchounikine 2002). Le champ de recherche concernant l'ingénierie EIAH est ainsi transdisciplinaire et les travaux nécessitent d'être menés par des équipes pluridisciplinaires.

Les sciences et technologies de l'information et de la communication (STIC), par leur objet même, ont de façon naturelle une contribution particulière à apporter aux dispositifs d'éducation et de formation et plus universellement à l'apprentissage humain. Cette contribution comprend d'une part des outils et des infrastructures favorisant la diffusion et la mutualisation de connaissances, et d'autre part la conception et la réalisation d'environnements informatiques dont la finalité explicite est de susciter et d'accompagner l'apprentissage humain (Derycke, Chevrin et al. 2005; Derycke, Y. Peter et al. 2005; Derycke, Hoogstoel et al. 2006). La recherche sur les EIAH « environnements informatiques pour l'apprentissage humain » est fortement pluridisciplinaire (elle nécessite en effet des points de vues croisés des domaines de recherches suivants : didactique des sciences, psychologie cognitive, informatique et au sein de l'informatique entre IHM, IA, génie logiciel, etc.) et doit donc s'instancier dans un tel cadre. Parmi les enjeux principaux des EIAH de nos jours, on peut citer les points suivants :

- L'étude et la formalisation du savoir-faire des enseignants comme fondements à la conception des EIAH et à la diffusion des connaissances,
- La prise en compte de situations concrètes d'apprentissage permettant de créer des cours évaluables, de valider les approches théoriques et pratiques, puis ensuite de généraliser,
- Mise au point de modèles d'évaluation adaptés à des EIAH, L'évaluation de la qualité de l'apprentissage de deux points de vue :
- Celui de la didactique des sciences : en quoi la mise en place de nouveaux modes de formations hybrides (présentiel/à distance) se traduit-il effectivement par l'apprentissage des connaissances attendues et par des compétences professionnelles pour les apprenants (ici de nouveaux professeurs)
- Celui de la psychologie cognitive : en s'appuyant sur des modèles théoriques récents, notamment en distinguant la mémorisation d'informations de l'habilité des utilisateurs à utiliser ces connaissances dans de nouvelles situations (transfert de connaissances). Au regard des travaux déjà conduits et des données de la littérature, il s'avère particulièrement crucial de disposer de données relatives aux processus effectivement mobilisés par les apprenants lors de leurs interactions avec des dispositifs de type EIAH afin de les mettre en relation avec les performances cognitives qu'ils sont aptes à produire ensuite,
- Le passage d'EIAH clos et locaux à des services Web adaptatifs/personnalisables accessibles de tous et de partout assurant : la réutilisation, le partage et l'échange des ressources d'apprentissages, et des connaissances et enfin le savoir-faire des enseignants comme fondements des EIAH, mais aussi pour assurer une meilleure formation des futurs enseignants et formateurs.

Cette recherche concerne plus particulièrement la mise en place de séquences de formation, à destination de futurs enseignants (PLC2 et PE), (appelés ci-après apprenants) basées sur l'utilisation de ressources multimédias et s'appuyant sur des savoir-faire d'experts et la modélisation de pratiques réelles. Il s'agit de mettre en place des scénarii didactiques qui permettent à de futurs enseignants « d'apprendre leur métier » en réalisant des séquences d'enseignement à partir d'une approche de type pédagogie active. La conception d'un EIAH nécessite de mettre en œuvre des théories didactiques, des théories de l'activité humaine, ainsi que des modèles et théories informatiques. Se pose alors la question de l'articulation entre la modélisation des situations et leur transposition informatique. Pour cette première étape, nous avons mis en œuvre une méthodologie de co-conception de scénarii didactiques fondée sur des pratiques réelles et une approche transdisciplinaire comme fondement d'un EIAH.

Ce type d'EIAH peut prendre plusieurs aspects : formation en autonomie, formation tutorée en présentiel / à distance, mise à disposition d'une documentation ou accès à des bases de données. Le projet porte plus particulièrement sur :

- Des apprenants en première et en seconde année d'IUFM : Professeurs des Ecoles (PE1) et Professeurs des Lycée et Collèges (PLC2) SVT et Physique-Chimie ;
- Deux formateurs SVT et trois formateurs Physique-Chimie que l'on considère comme experts ;
- Une thématique commune (déclinée dans les programmes officiels de l'Ecole et de seconde généralisée) : « L'air en tant que gaz dans ses aspects statiques et dynamiques : propriétés, théorie et applications ».

Pour créer un EIAH dans le contexte des CAREST décrit ci-dessus, les contraintes sont les suivantes :

- Il ne suffit pas de rendre accessible aux apprenants un environnement et des documents, il y a nécessité d'un formateur avec une intention didactique qui organise le temps, l'espace, et le milieu d'apprentissage (Tricot and Plegat-Soutjis 2003). Celui-ci définit et régule ensuite l'activité.
- La mise en place des scénarii de formation dépend de variables et de paramètres dont
- L'origine des apprenants (ici PE et PLC en Formation Initiale (FI) avec des variabilités interindividuelles et inter catégorielles,
- La situation didactique,
- Les ressources humaines et documentaires disponibles,
- La répartition entre formation en présentiel et à distance
- La distribution de l'activité entre les différents « participants » (machine, apprenants, tuteurs, enseignants, etc.).

Ces variables et paramètres permettront de déterminer des stratégies d'adaptation pour l'EIAH. En d'autres termes, il s'agit de définir un scénario générique qui sera à même de couvrir une majorité de situations didactiques. A partir de ce scénario générique, l'EIAH calculera dynamiquement un scénario dédié à l'apprenant courant et sa situation didactique. Les enseignants appartiennent à une communauté de pratiques : ils partagent des connaissances communes pour travailler ensemble. Le scénario générique est basé sur les connaissances communes explicites qui sont les points clé de la conception de l'EIAH et de l'adaptation pour fournir à chaque apprenant un cours adapté à sa situation didactique.

Dans un premier temps, nous définissons notre projet de recherche, ses objectifs et les contraintes de construction de l'EIAH à partir de la grille d'analyse de Platon-1 (Tchounikine and Al. 2004). A partir de cette analyse, les principes fondamentaux de conception de l'EIAH d'un point de vue informatique sont présentés. Dans un troisième temps, nous explicitons le plan de travail du projet sur deux ans, la méthode d'acquisition des scénarii didactiques fondée sur la théorie anthropologique didactique du savoir et la théorie de l'activité, ainsi qu'une première modélisation informatique fondée sur les modèles de tâches hiérarchiques de l'intelligence artificielle et un modèle de pédagogie active. Les scénarii didactiques et la modélisation informatique font l'objet d'une analyse croisée pour articuler ces outils conceptuels et les modélisations formelles de l'informatique pour la conception du dispositif de formation. En conclusion, une première discussion sur les apports de chaque théorie ou modèle est abordée, ainsi que les perspectives envisagées.

5.3 Problématiques et méthodologie

5.3.1 Principes fondamentaux de conception de l'EIAH

L'EIAH doit être accessible sur Internet, capable d'engendrer dynamiquement des artefacts avec une flexibilité maximale, adaptatif, capable d'assurer la réutilisation, le partage et l'échange des ressources d'apprentissages, des connaissances et du savoir-faire des enseignants. Il est au moins nécessaire d'assurer l'interopérabilité au niveau des ressources. La génération dynamique, l'adaptation et la réutilisation des ressources, peut être réalisée d'une part grâce aux documents virtuels adaptatifs, notés DVA et d'autre part grâce au web sémantique.

Le moteur de composition flexible SCARCE - SemantiC and Adaptive Retrieval and Composition Engine – servira d'environnement de base et sera étendu pour satisfaire aux besoins du projet. Avec cet environnement, la conception d'un EIAH est centrée sur les connaissances explicites d'une communauté de pratiques : connaissances et savoir faire des enseignants. Ces dernières permettent de formaliser les différents modèles sous-jacents d'un l'EIAH et donc de spécifier la sélection, le filtrage et l'organisation du moteur de composition – le scénario générique dans notre cas. Dans le contexte du projet MODALES, ce sont les modèles de scénarii munis de leurs variables et paramètres qui permettront de définir et de formaliser la spécification de l'EIAH. L'adaptation peut être analysée à l'aide des catégories d'adaptation suivantes : variabilité inter catégorielles, variabilité intra catégorielle, présentiel ou à distance, activité machine ou humaine et type de ressources documentaires. Ces catégories mèneront à des objectifs et des stratégies d'adaptation différents. Le moteur de composition SCARCE utilise quatre catégories de modèles qui sont formalisés par des ontologies :

- Le schéma de métadonnées qui décrit la structure d'indexation des ressources dont certains domaines de valeurs sont pris dans le modèle de domaine et le modèle de document,
- Un modèle de domaine représentant des connaissances relatives à la physique, la didactique et l'épistémologie,
- Un modèle de document et un modèle d'adaptation qui sont fortement couplés
- Un modèle utilisateur qui représente les catégories d'apprenants et d'enseignants, ainsi que des propriétés individuelles. Les scénarii permettront de déterminer les modèles de document et d'adaptation, le schéma de métadonnées et le scénario générique ou spécification.

5.3.2 Questions de recherche

Les objectifs de la recherche présente une part appliquée importante puisque que nous cherchons à :

- Améliorer le travail des formateurs par la mise à disposition de cet EIAH disposant de scénarii didactiques fondés sur des pratiques, muni de nombreux paramètres pour les adapter aux apprenants visés et d'une base de ressources pédagogiques notionnelles, didactiques et épistémologiques/historiques) et par la mise au point d'une méthodologie de conception de module d'autoformation ;
- Améliorer le travail des apprenants de l'IUFM car il s'agit de mettre en ligne un EIAH proposant des situations de formations innovantes, dans une approche par pédagogie active, situations associées à une base de données de ressources documentaires riches notionnelles, didactiques et épistémologiques/historiques) afin de les aider à fabriquer des séquences d'apprentissage en sciences ;

Il présentent aussi une part théorique, puisqu'il s'agit de produire des résultats sur des problèmes de didactique des sciences et de formation des maîtres par des travaux focalisés sur le processus didactique dont le point d'entrée est la notion d'enjeu de connaissances et de modélisation de la situation qui en permet l'acquisition.

Ces travaux se justifient en eux-mêmes mais sont aussi nécessaires à l'élaboration de l'EIAH qui se doit d'être transdisciplinaire.

On peut notamment citer les problématiques suivantes :

Dans le champ de la modélisation des situations didactiques pour générer l'EIAH :

- Q1 : Quels sont les modèles d'enseignement/apprentissage réels sous-jacents aux pratiques des formateurs ? Quels sont les modèles épistémologiques de références des formateurs ? Comment ces modèles influent-ils l'élaboration des dispositifs didactiques ainsi que l'utilisation des systèmes ?

- Q2 : En quoi le rôle du professeur est-il modifié dans une distribution à distance/présentiel ? Comment, en l'absence du professeur, sont transmises ses attentes qui sont d'habitude signifiées par des éléments non-verbaux ? Le professeur anticipe-t-il ces questions, et si oui, de quelle manière lorsqu'il met en place les parties distantes de son scénario ?
- Q3 : Quelle est la répartition des tâches dans le dispositif didactique entre l'homme et la machine ? Quels sont les gestes professorales susceptibles de supporter une modélisation formelle informatique ? Quelles sont les théories de références pertinentes pour la mise au point d'un modèle ?
- Q4 : Quel est le rôle de la documentation (en ligne) ? En particulier, celle concernant la didactique, l'épistémologie et l'histoire des sciences ? Quels sont les effets attendus par les formateurs et en quoi les pratiques professionnelles des PE et PLC sont-elles modifiées ? Quels éléments d'une transposition didactique externe et interne peut-on mettre en évidence ?

Dans le champ de l'évaluation des apprentissages :

- Q5 : Quelle méthode pour évaluer les effets de l'EIAH en terme d'apprentissage ? Au niveau de la psychologie cognitive, nous nous interrogeons notamment sur l'effet des améliorations ergonomiques de l'EIAH et sur les effets de ces améliorations sur les traitements cognitifs réalisées pendant les phases d'apprentissages. Autrement dit, il s'agira d'évaluer en quoi la manière dont l'information est présentée dans l'EIAH va modifier ou non l'interaction et influencer son produit : l'apprentissage. Un autre champ d'étude à développer concerne les approches plus dynamiques de l'apprentissage à distance où l'instruction peut être adaptée en temps réel au niveau d'expertise des apprenants (adaptative e-learning). Au niveau de la didactique des sciences, il s'agit d'étudier en quoi la mise à disposition d'une base de données dans trois domaines (notions, informations didactiques, épistémologie/histoire des sciences) associée à un type de scénario de formation professionnelle (adapté pour un profil novice/expert) correspond aux effets attendus en terme de compétences professionnelles.

Du point de vue de l'informatique, il s'agit : i) de concevoir un EIAH capable d'engendrer dynamiquement des artefacts avec une flexibilité maximale. La mise en place des scénarii de formation dépend de paramètres assez nombreux qui ont pour effet d'accroître le nombre de versions d'EIAH. Il devient donc nécessaire d'automatiser leur génération ii) de passer d'EIAH clos et locaux à des services et/ou sites Web adaptatifs/personnalisables accessibles à tous et de partout assurant : la réutilisation, le partage et l'échange des ressources d'apprentissages, notamment les connaissances et savoir-faire des enseignants. Ces dernières permettront de plus une meilleure formation des futurs enseignants et formateurs puisqu'elles sont explicitées et formalisées, donc plus facilement transmissibles.

L'EIAH doit servir de support à la compréhension et permettre de tester des hypothèses (par la mise en place d'un dispositif expérimental) concernant : a) la modélisation des activités des apprenants et des formateurs, b) les effets sur les pratiques des apprenants de la mise à disposition d'une documentation de type didactique et historique et de la pluralité des modalités de scénarii sous-tendus par des épistémologies différentes. Cela devrait se

traduire dans les classes par une modification des pratiques professorales jusque là souvent inductivistes.

Les finalités de l'artefact informatique sont de : i) disposer d'un EIAH pour un groupe restreint, les acteurs du CAREST, ii) disposer de retours d'usages ou de données pour contribuer à la compréhension, iii) obtenir une mise à l'épreuve/validation de scénarii didactiques pour un apprentissage mixte présentiel/distance pour l'enseignement des sciences, iv) identifier des obstacles à l'apprentissage, déterminer la corrélation entre les activités des apprenants au sein de l'EIAH et la modification dans leur pratique professionnelle, v) faire évoluer les pratiques professionnelles des apprenants (ici de jeunes enseignants), changer les méthodes de conception de cours, les fonder sur des théories didactiques et les mettre en pratique plus tard, l'apprenant étant un futur enseignant, mais aussi évaluer l'impact du notionnel (ici domaine de la physique), de la didactique et de l'histoire des sciences sur les pratiques de conception de cours pour les futurs enseignants, vi) concernant l'ergonomie de l'interface, une réflexion est à mener sur les formats de présentation de l'EIAH, en visant à réduire la charge cognitive inutile en adoptant un mode de présentation de l'information pertinent pour la tâche en cours.

5.4 Méthodologie de co-conception des scénarii didactiques

Nous suivons la démarche indiquée lors du dépôt du projet DALE :

- Etape 1 : élaboration par les formateurs experts des scénarios d'apprentissage ;
- Etape 2 : choix des outils conceptuels et explicitation de leurs apports, issus de la théorie de l'activité (Cole and Engeström 1993), de la théorie des situations didactiques (Brousseau 1998) et de l'action du professeur (Sensevy, Mercier et al. 2000), de la théorie anthropologique didactique du savoir (Chevallard 1992; Chevallard 1999) ;
- Etape 3 : Modélisation des scénarios, articulation entre ces outils conceptuels de modélisation des pratiques réelles et les modélisations formelles de l'informatique pour la conception d'un dispositif de formation ;
- Etape 4 : élaborer une première plate-forme d'autoformation expérimentale non adaptative afin de valider les scénarii proposés, leurs paramètres et l'impact en termes d'apprentissage.
- Etape 5 : analyser les activités des apprenants et des formateurs et évaluer les effets de ce premier dispositif non adaptatif. Il s'agit notamment d'évaluer et consolider les scénarii et d'acquérir les paramètres effectifs de l'adaptation.
- Etape 6 : élaborer la seconde plate-forme adaptative à partir des scénarii consolidés et des paramètres associés (pertinents). Il s'agit ensuite d'évaluer notamment en termes d'apprentissage cet environnement adaptatif.

Les étapes 4 et 5 sont en cours et constitueront avec l'étape 6 la seconde phase du projet.

Le processus de co-conception peut être vu comme un modèle tourbillonnaire muni de boucles successives qui favorisent la coopération entre les informaticiens et les enseignants experts à l'aide d'interactions réitérées - semblables à la proposition faite par Akrich et al 1988 (Akrich, M. Callob et al. 1998). Avec une telle approche, plusieurs modèles et artefacts (environnement informatique) sont réalisés. Plusieurs théories seront employées pour acquérir les pratiques des enseignants et pour déterminer les avantages et les limites de chaque théorie. Ces dernières sont : théorie de l'activité (Cole and Engeström 1993), la théorie des situations didactiques (Brousseau 1998) et la théorie en anthropologie didactique de la connaissance (Chevallard 1992; Chevallard 1999). Actuellement, la conception des scénarios est fondée sur des pratiques réelles acquises à l'aide de la théorie en anthropologie didactique de la connaissance et est formalisée dans un modèle de tâches hiérarchiques (Wielinga, Velde et al. 1992; Willamowski 1992; Duursma, Olsson et al. 1993; Trichet 1998). Les enseignants appartiennent à une communauté de pratiques : ils partagent des connaissances communes pour travailler ensemble (Wenger 1998). Les scénarios sont fondés sur les connaissances explicites de la communauté de pratiques.

Le schéma de la transposition didactique de Perrenoud (Perrenoud 1998) résume notre méthodologie pour concevoir et modéliser les situations d'apprentissage : i) repérer et décrire finement les pratiques et le savoir-faire des professeurs et des étudiants, ii) identifier les compétences à l'œuvre (des professeurs et des étudiants), iii) analyser les ressources cognitives (la connaissance, etc.) utilisé par des professeurs et des étudiants, iv) faire des hypothèses quant au mode de genèse des compétences en situation de formation (pour les étudiants) ; v) élaborer des dispositifs, des situations, des contenus planifiés de la formation (un curriculum formel) et les mettre en application (un curriculum réel). La conception a été faite en trois étapes principales : 1) la conception d'une première version des scénarios par les professeurs experts, 2) acquisition des scénarios raffinés employant une théorie en anthropologie didactique de la connaissance (Chevallard 1992; Chevallard 1999) 3) Formalisation des scénarios dans un modèle de tâches hiérarchiques (Wielinga, Velde et al. 1992; Willamowski 1992; Duursma, Olsson et al. 1993; Trichet 1998). L'étape 1 a été faite pour initialiser le modèle tourbillonnaire et la coopération entre les enseignants experts et les informaticiens. Les étapes 2 et 3 mènent aux boucles itératives étudiant en détail la théorie, les modèles et les artefacts. Ce paragraphe est organisé comme suit : tout d'abord, nous décrivons la première version des scénarios conçus par les professeurs experts ; deuxièmement, nous présentons la théorie de Chevallard et l'amélioration correspondante des scénarios et troisièmement nous présentons quelques propriétés des modèles de tâches hiérarchiques et leurs relations avec la théorie de Chevallard.

5.5 Elaboration des scénarios d'apprentissage

Les étudiants sont de futurs enseignants : instituteurs (appelés PE pour " professeur des Ecoles") et les professeurs d'école secondaire (appelés PLC pour "professeurs des Lycées et Collèges " : sciences et vie de la terre et physique). Les professeurs sont considérés comme experts en matière d'éducation. L'objet du cours est « l'air comme gaz dans ses aspects statiques et dynamiques : propriétés, théorie et applications » pour différentes catégories de professeurs. Un scénario d'étude commun PO (dont les variables sont les étudiants, le professeur expert et les ressources disponibles) a été établi (cf. Tableau 3). Il montre deux phases : 1) construction des références professionnelles pour enseigner, 2) développement d'une séquence de formation mise en œuvre dans des salles de classe. Une étape se compose

d'une séquence d'activités. Les scénarios pour PE et PLC en sciences et vie de la terre et en physique ont été établis selon la même procédure : chaque professeur expert (selon son domaine) a établi un scénario du plan commun PO. Il associe à chaque activité des deux phases des informations pertinentes : 1) activation de la phase (si elle existe dans le scénario) ; 2) distance ou face à face ; 3) une description des ressources disponibles et de leur type (physique, didactique, épistémologique/historique) ; 4) une description d'activité pour des étudiants et des professeurs. Les activités Tij peuvent être raffinées en fonction de l'étudiant, de son groupe, de la situation didactique et quelques autres contraintes données par le professeur.

Scénario MODALES : phase 1	Scénario MODALES : phase 2
Construction de références professionnelles	Elaboration d'une séquence d'apprentissage
<u>Etapas</u>	<i>Etapas</i>
T11. Construction de la problématique scientifique	T21. Construction du problème didactique
T12. Lecture de la documentation initiale	T22. lecture de la documentation
T13. Démarche explicative (formulation d'hypothèses, de procédures expérimentales avec prédictions des résultats)	T23. Définition des objectifs scientifiques visés
T14. Ecrit intermédiaire	T24. Objectifs de méthodes et de savoir-faire
T15. Mise en œuvre et réalisation expérimentale	T25. Description de la procédure de résolution du problème didactique
T16. Production d'un écrit : compte-rendu d'expériences	T26. Descriptif des activités à faire
T17. Confrontation des comptes-rendus (Forum)	T27. Confrontation des productions (Forum)
T18. Synthèse et apports théoriques : notions, épistémologie, didactique, histoire des sciences, méthodes	T28. Synthèse et validation des procédures

Tableau 3 : Plan de scénario de formation

A titre d'exemple, nous illustrons le résultat de cette élaboration (phase 1) pour la formation des PE (professeurs des Ecoles) :

- T11 : Consigne donnée par le formateur : « Problème : La matérialité de l'air. Proposer deux expériences pour prouver l'existence de l'air. », (A distance)
- T12 : Mise à disposition d'une documentation en ligne essentiellement notionnelle. (A distance)
- T13/T14: A partir de la lecture de la documentation, les PE en binôme rédigent un écrit intermédiaire et l'envoient au formateur (à distance)
- T15 : Les PE montrent leur système de preuves avec le matériel du CAREST. Le formateur aide et valide les démarches. (en présentiel)
- T16 : un compte-rendu définitif est rédigé par chaque binôme. Il est communiqué à tous. (à distance)
- T17 : Forum de discussion sur la question : qu'est qu'une preuve en sciences » (PE+formateur : à distance)

- T18 : Synthèse de la séquence et apports universitaires dans le domaine de la physique (l'air), l'épistémologie/Histoire des Sciences, la didactique des sciences (textes officiels, obstacles, ressources, etc.). Mise à disposition de la base de données correspondante. (en présentiel)

La seconde phase a ensuite pour objet de résoudre le problème didactique : « élaborer une séquence d'apprentissage en CIII sur la matérialité de l'air ».

Trois scénarios (PE, PLC Physique et SVT) ont donc été ainsi élaborés. Ils présentent tous des spécificités en fonction de plusieurs variables dont les types d'apprenants (PE/PLC Physique/PLC SVT ; novice/expert dans le domaine), l'interaction apprenant/formateur (à distance/e présentiel), la documentation mise à disposition. A partir de ces données, une analyse approfondie des scénarii experts a été menée à partir de la théorie praxéologique de Chevallard.

5.5.1 Analyse des scénarios

La théorie de Chevallard a été utilisée pour raffiner l'acquisition des pratiques réelles des professeurs experts. Selon Chevallard (Chevallard 1992), l'action du professeur peut s'exprimer en termes de types de tâches T accomplies au moyen d'une technique τ . Cette association tâche-technique définit un savoir-faire qui s'appuie sur un environnement technologico-théorique (ou savoir) formé d'une technologie θ (discours qui justifie et rend intelligible la technique) et une théorie Θ justifiant et éclairant cette technologie. Ce système à 4 composantes ($T/\tau/\theta/\Theta$) constitue une organisation praxéologique articulant une partie pratico-technique (savoir-faire) et une partie technologico-théorique (savoir). Toute praxéologie est donc constituée par un système de tâches autour duquel se développent et s'organisent techniques, technologies et théories (Laubé, Garlatti et al. 2005; Laubé, Garlatti et al. 2005; Laubé, Garlatti et al. 2006; Laubé, Garlatti et al. 2006; Laubé, Kuster et al. 2006). Dans le cas qui nous intéresse, celui de la formation des maîtres, on voit que ce point de vue peut s'appliquer à deux niveaux :

- Le formateur expert qui possède son système ($T/\tau/\theta/\Theta$)formateur que nous cherchons à mettre en évidence puisque nous considérons que ce système constitue la référence pour la conception d'un EIAH adaptatif,
- L'apprenant qui est en formation à l'IUFM. En effet la procédure d'apprentissage (les activités) mise en place par le formateur dans le cadre de son système ($T/\tau/\theta/\Theta$)formateur ont pour objectif de faire évoluer le système de l'apprenant ($T/\tau/\theta/\Theta$)apprenant d'un état initial vers un état final.

On retrouvera différents moments de l'organisation didactique (Chevallard 1999) :

- Le moment de la première rencontre avec le type de tâches $T_{\text{apprenant}}$; (M1)
- Le moment de l'exploration du type de tâches $T_{\text{apprenant}}$ et de l'élaboration d'une technique ; (M2)
- Le moment de l'élaboration technologico-théorique relatif à une technique (M3)

- Le moment du travail de la technique qui doit améliorer la technique en la rendant plus efficace et plus fiable et accroître la maîtrise que l'on en a. (M4)
- Le moment de l'institutionnalisation (M5)
- Le moment de l'évaluation qui s'articule au moment de l'institutionnalisation. (M6)

Il est pertinent d'analyser un scénario selon l'approche de Chevallard. Illustrons le point de vue des apprenants à partir de la phase 1 du scénario expert (construction de références professionnelles) à destination des PE néophytes.

Celle-ci a pour objectifs de :

- Traiter la thématique «La matérialité de l'air» par résolution de problème («Comment prouver que l'air existe? Proposer deux expériences.»);
- Apprendre une technique τ de résolution de problèmes (parmi d'autres techniques),
- Analyser la démarche d'investigation mise en œuvre.

Les œuvres étudiées sont ici «la théorie de la matière» et «la démarche en sciences à l'Ecole» (donc la didactique des sciences et l'épistémologie). La technique τ de résolution du problème est guidée par le formateur et passe par trois sous-tâches principales T (les M1, M2, etc. font référence aux moments de l'organisation didactique) :

- (T₁) Proposition écrite de résolution validée par le formateur (M1);
- (T₂) Expérimentation, analyse et validation des solutions individuelles (M2, M3);
- (T₃) Institutionnalisation des aspects théoriques, épistémologiques et didactiques à partir de l'analyse des productions des apprenants et de la technique τ suivie.

Cette institutionnalisation (M5) s'appuie sur une documentation référence en physique, en didactique et en épistémologie/histoire des sciences (élaboration d'un savoir technologico-théorique). On trouve dans la phase 2 du scénario expert «Elaboration d'une séquence d'apprentissage» (voir annexe I), de façon implicite, le moment de l'évaluation (M6), car il est attendu des apprenants une transposition didactique de la phase 1. Lors de la session de formation suivante (et en fonction de l'évaluation), le scénario sera proposé avec la phase 1 du scénario expert «Construction de références professionnelles» (voir Annexe I) où la situation didactique constituera le moment du travail de la technique (M4).

Chacune des trois sous-tâches T repose sur un ensemble technique espéré être routinier (échanges de mails intra-groupe et avec le formateur, recherches documentaires, etc.) qui peut constituer un obstacle dans le cas contraire. Les informations disponibles dans les scénarii experts ne permettent pas de décrire l'organisation de ces tâches entre-elles. Celle-ci sera mise en évidence lors de l'observation des usages.

De même nous pouvons analyser l'activité du formateur et en déduire en partie dans cette situation la partie pratico-technique de son système $(T/\tau/\theta/\Theta)_{\text{formateur}}$. Cependant, seul un

entretien d'explicitation avec les formateurs permet d'accéder à la partie technologico-théorique. Pour ce faire, nous avons filmé les séances de négociation et l'élaboration par les formateurs du plan de scénario de formation commun **P₀**. Quatre thèmes, qui ont fait l'objet de débats entre les formateurs, ont ainsi été repérés : leur conception de la science, leur représentation des élèves professeurs, leur représentation de la formation à donner aux élèves de lycée, la représentation du métier et de la formation à dispenser aux futurs enseignants. A partir de ces quatre catégories, nous avons mené une série d'entretien (de 75mn) avec chaque formateur suivant un protocole commun. Leur analyse (qui est en cours) va permettre de mettre en évidence les facteurs qui interviennent dans les choix didactiques des formateurs quand ils proposent des activités et de faire expliciter les théories sous-jacentes et souvent implicites qui motivent leurs choix. Nous mettrons en évidence les différences qui peuvent apparaître suivant la discipline d'origine (Physique-chimie ou SVT) et le public ciblé (PE ou PLC).

Un scénario se compose de deux parties principales : la première est consacrée à la "construction des références professionnelles du professeur" et la seconde à la "construction d'une leçon de salle de classe" (cf. Figure 49). L'analyse des deux phases du scénario au moyen du système (T/t/θ/Θ) et des différentes parties de l'organisation didactique a apporté les résultats suivants :

- On peut mettre en évidence une même structure de formation dans chacune des phases : a) proposition d'un problème à résoudre (T11 et T21), b) construction par les apprenants d'un système Tâches/Technique pour résoudre le problème (T12 à T16 et T22 à T26), c) construction (au sein d'un forum) d'un discours critique sur les systèmes Tâches/Techniques : construction d'une technologie (T17 et T27), d) institutionnalisation par le formateur qui apporte un discours théorique qui valide la technologie (T18 et T28)

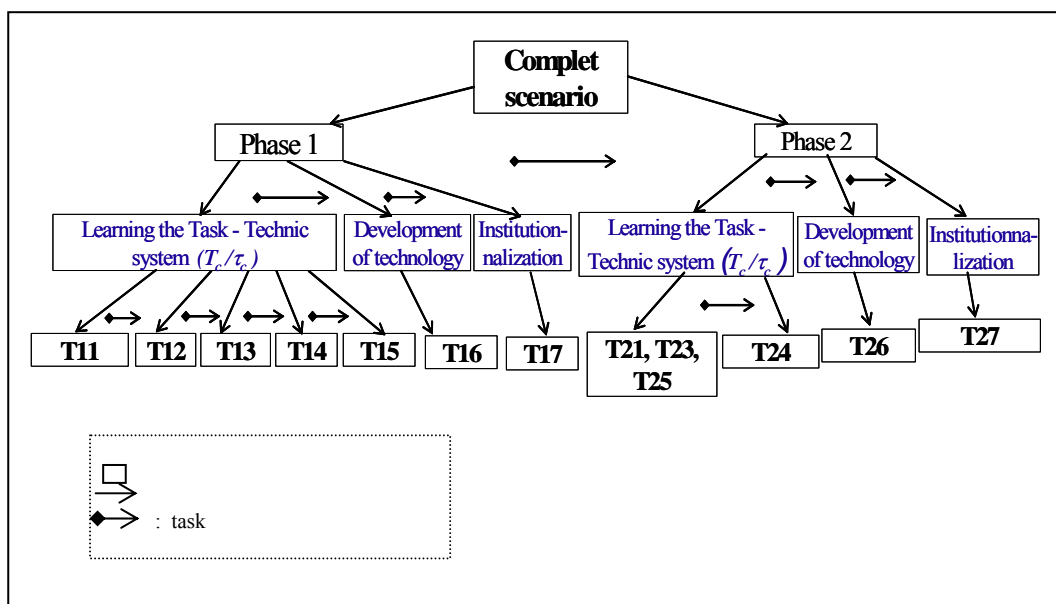


Figure 49 : Une organisation des étapes du scénario P0.

- On peut observer plusieurs types de système tâches/technique caractérisés par le rôle du formateur :
 - Routinier : on n'observe aucune intervention du formateur et celui-ci n'a pas prévu d'intervenir (par exemple, les apprenant sont considérés comme sachant lire et écrire) sauf éventuellement pour donner le contrat de tâche ainsi que les ressources et contrôler que la tâche a été effectuée en temps et en heure. Ce contrôle s'effectue sans problème à distance.

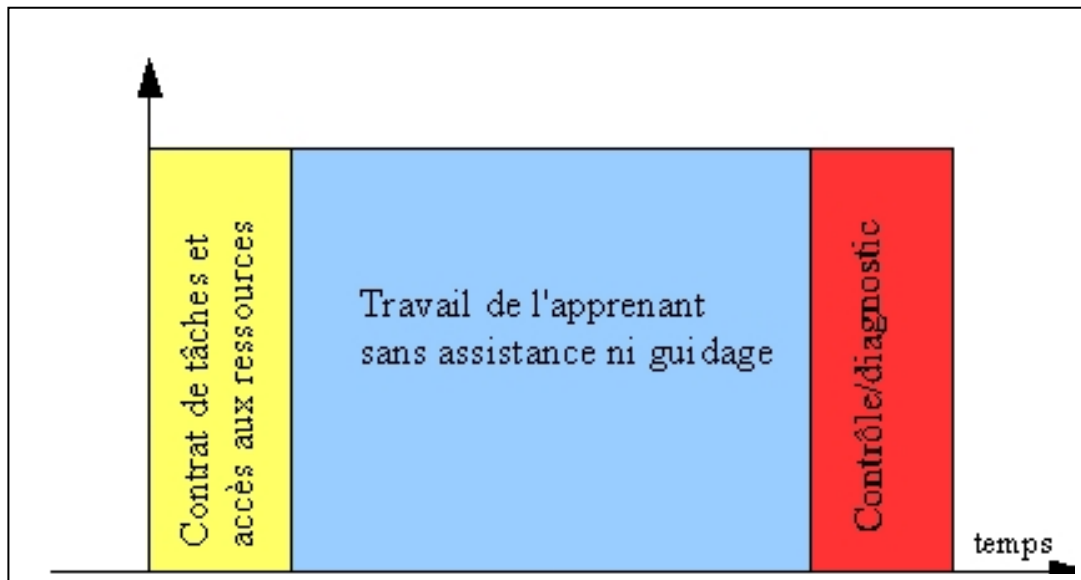


Figure 50 : Schéma d'un système tâche/technique routinier

- La figure 50 montre trois moments : le premier moment $[t_0, t_\alpha]$ est celui où le formateur qui a défini le milieu didactique, indique la tâche à effectuer, la date pour le rendu. Le second $[t_\alpha, t_\beta]$ correspond au temps imparti pour effectuer la tâche sans aide ni assistance de la part du formateur. La technique est laissée au libre choix de l'apprenant. Le dernier moment $[t_\beta, t_f]$ est celui du contrôle que la tâche a bien été effectué à l'heure dite. Les premiers et derniers moments peuvent être totalement absents, le grain le plus fin de la description sera alors choisi par le formateur et décrira une tâche/technique de l'apprenant.
- En apprentissage : ici, le professeur intervient dans la situation didactique après une durée $\Delta\tau$. La valeur de ce curseur $\Delta\tau$ donne plusieurs indications. Plus le professeur intervient rapidement dans la situation didactique (i.e. $\Delta\tau$ est petit), plus le système tâches/techniques est considéré comme problématique. On peut mettre ainsi en évidence une gamme de situations didactique : 1) premier contact avec le problème et donc première élaboration d'un système tâches/technique pour résoudre le problème, 2) travail d'un système tâches/technique problématique déjà éprouvé à la suite du premier contact, 3) travail d'un nouveau système de tâches/technique plus performant, mais problématique.

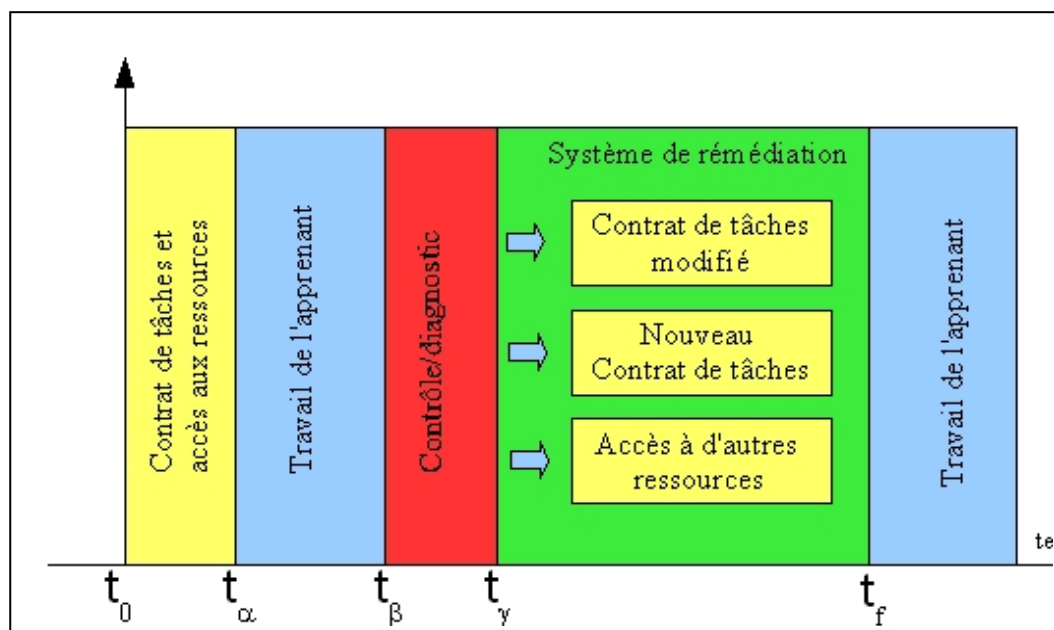


Figure 51: Schéma d'un système tâche/technique en apprentissage

La figure 51 montre quatre moments tous présents dans le scénario :

- Le premier moment $[t_0, t_\alpha]$ est celui où le formateur qui a défini le milieu didactique, indique la tâche à effectuer, les consignes à suivre, la date pour le rendu, les ressources à disposition, etc. Dans ce moment, on peut (ou non) donner la possibilité de questionner le formateur
- Le second moment $[t_\alpha, t_\beta]$ correspond au travail de l'apprenant. Une aide ou une assistance peut être demandée de la part de l'apprenant. Cette aide amène à un système de remédiation en fonction du diagnostic fourni par le formateur (ou la machine) et qui renvoie au moment suivant. La technique peut être imposée ou au libre choix de l'apprenant en fonction de la situation didactique en jeu (premier contact, travail d'une technique, etc.)
- Le troisième moment $[t_\alpha, t_\gamma]$ constitue celui du contrôle et du diagnostic. Ici va être effectuée une analyse de la situation qui est identifiée comme problématique pour l'apprenant. En fonction du mode présentiel/à distance, le contrôle peut être très diversifié : de l'ordre de la communication non verbale (en présentiel) au dépôt d'un document en ligne (à distance). L'analyse de la situation didactique par le formateur expert le conduira à choisir en fonction des difficultés parmi un tableau de remédiations.
- Le quatrième et dernier moment $[t_\gamma, t_f]$: les remédiations. Elles peuvent constituer un simple aménagement du contrat de tâches initial à la proposition d'un autre et nouveau contrat qui se traduit par une nouvelle phase de travail de l'apprenant.

Pour en terminer avec ce point, si nous reprenons la taxonomie de Pernin et Le Jeune (Pernin and Lejeune 2004), nous possédons ainsi les outils pour formaliser un scénario

descriptif, adaptable en vue de la conception d'un scénario automatisable. La granularité la plus fine de formalisation (l'activité élémentaire la plus petite) s'appuie ici nécessairement sur un des deux systèmes tâches/technique décrit ci-dessus (routinier et en apprentissage). Reste un paramètre qui jouera un rôle important quant à la répartition des rôles entre formateur, apprenants et machine. L'expertise du formateur concernant le public (novice/expert), le système de tâches/techniques mise en apprentissage induit une situation didactique et une répartition en présentiel/à distance : beaucoup de présentiel s'il s'agit d'un public novice en premier contact avec un problème : des PE en début de formation, beaucoup plus de distance s'il s'agit d'un public déjà expert : des PE en fin de formation. Cela peut aller jusqu'à la réduction de la phase 1 à une simple mise à disposition de ressources notionnelles, didactiques et épistémologiques/historiques pour introduire la phase 2 « construire une séquence d'apprentissage »

5.5.2 Formalisation des scénarios par un modèle de tâches hiérarchiques

Les modèles hiérarchiques de tâches représentent des connaissances de résolution des problèmes. De nombreuses études de ces modèles ont été réalisées (cf. annexe 1 pour un état de l'art). Ces modèles peuvent être caractérisés par de nombreuses propriétés pour décrire ces connaissances de résolution des problèmes. Dans notre cas, nous sommes seulement intéressés par ceux qui sont pertinents pour formaliser les scénarios. Nous devons donc analyser les concepts de tâches (dénotté en italique gras pour les distinguer des tâches dans la théorie anthropologique didactique de la connaissance dénotées T_c), de méthode, de tâche abstraite et élémentaire, de structure de contrôle des tâches et sous tâches, graphe de composition et d'héritage des tâches (Garlatti, Tetchueng et al. 2006).

Dans le cadre du paradigme Tâches/méthodes des modèles hiérarchiques de tâches, les tâches définissent des activités et des sous activités contrôlées par un système des gestion de connaissance (Trichet 1998). Il existe deux types de tâches : les tâches abstraites et élémentaires. Une tâche abstraite représente une activité de haut niveau qui se compose de sous tâches. Les sous tâches peuvent être abstraites ou élémentaires. Une tâche élémentaire ne se compose pas de sous tâches. Elle peut être réalisée par une simple procédure - par exemple, une procédure de recherche documentaire, une interaction homme/machine particulière, etc... Ainsi, Une tâche abstraite peut être décomposée récursivement en sous tâches jusqu'à l'obtention de tâches élémentaires. Une méthode décrit comment une tâche particulière peut être réalisée. Les méthodes définissent la structure de contrôle qui permet la décomposition récursive des tâches en sous tâches et la structure de contrôle qui définit l'ordre d'exécution des sous tâches. Pour une tâche donnée, plusieurs méthodes peuvent être employées pour la réaliser. Dans ce cas-ci, un mécanisme de sélection choisit dynamiquement la méthode appropriée pour réaliser la tâche en fonction du contexte courant de résolution de problème. Par ailleurs, les tâches sont également organisées dans un graphe d'héritage permettant de raffiner les tâches en tâches plus spécifiques.

Une comparaison des concepts (et de leurs propriétés) de la théorie de Chevallard appliquée aux scénarios et ceux des modèles hiérarchiques de tâches montre des similitudes sémantiques entre elles. En effet, en fonction de leurs interprétations et propriétés respectives, nous pouvons établir les correspondances suivantes :

- Les tâches et les sous tâches T_c de la théorie de Chevallard, peuvent être représentés par le concept de Tâches du modèle hiérarchique de tâches.
- Les techniques qui sont une manière de réaliser une tâche T_c peuvent être représentées par le concept de méthode décrit ci-dessus.

Parmi les tâches T_c de la théorie de Chevallard, nous peut distinguer les tâches d'apprentissage correspondant à des paires problématiques et des tâches sans apprentissage correspondant à des paires routinières) ; Les premières sont généralement décomposées récursivement en sous tâches d'apprentissage et en tâches sans apprentissage. Les secondes ne peuvent pas être décomposées en tâches d'apprentissage charge. Nous pouvons ainsi représenter un graphe de composition de tâches d'apprentissage, dans lequel, les tâches d'apprentissage sont vues comme des tâches abstraites et les tâches sans apprentissage comme des tâches élémentaires.

- La décomposition d'une tâche T_c en sous tâches T'_c par une technique peut être représenté par le mécanisme de décomposition d'une tâche par une méthode dans le modèle hiérarchique de tâches.
- Les étapes T_{ij} du scénario (cf. Tableau 3) sont spécialisées selon les propriétés spécifiques de l'apprenant. Ceci peut être vu comme spécialisation de la définition de la tâche et est représenté dans un graphe d'héritage dans les modèles hiérarchiques de tâches.
- La liste ci-dessus montre que certains concepts et propriétés de la théorie de Chevallard peuvent être transposés dans des concepts et mécanismes des modèles hiérarchiques de tâches. Cette formalisation montre qu'il est possible de transposer dans un modèle hiérarchique de tâche la structure T_c/t_c de l'organisation de Chevallard de l'étudiant et du professeur. Les modèles hiérarchiques de tâches nous permettent de transposer la structure hiérarchique T_c/t_c de la théorie de Chevallard et de leurs propriétés dans un modèle sur ordinateur. Mais, il nous permettra aussi de mettre en commun des activités entre des acteurs et de contrôler les activités à distance et en présentiel.

Actuellement, il n'est pas encore possible de formaliser tous les aspects du savoir-faire des professeurs parce que, la description des scénarios n'est pas finie. En d'autres termes, le processus de co-conception doit se poursuivre en allant à la prochaine boucle de notre modèle tourbillonnaire.

5.6 Rôle des scénarios dans la conception de l'EIAH

Le système de E-learning adaptatif peut être vu comme document virtuel adaptatif. Il utilisera un moteur de composition flexible, appelé SCARCE et fondé sur l'approche du web sémantique (Garlatti and Iksal 2004; Garlatti, Iksal et al. 2004; Iksal and Garlatti 2004). Dans notre cadre, un document virtuel adaptatif se compose d'un ensemble de ressources, de leurs métadonnées correspondantes, de différentes ontologies et d'un moteur de composition adaptatif qui peut sélectionner les ressources pertinentes, les organiser et les assembler en adaptant le document fourni aux besoins des apprenants et à la situation d'apprentissage

courante. Pour garantir la flexibilité, la sélection, l'organisation et l'adaptation sont des paramètres du moteur de composition et peuvent être spécifiés. Cette spécification s'appelle un scénario générique. Il doit permettre de traiter la plupart des situations d'apprentissage. A partir d'un scénario générique, le système de E-learning calculera dynamiquement un scénario particulier dédié à l'apprenant courant et à sa situation d'apprentissage. SCARCE utilise quatre ontologies faiblement couplées qui sont : l'ontologie des métadonnées au niveau de l'information qui décrit la structure d'indexation des ressources, quelques valeurs d'indexés sont pris dans les ontologies de domaine et de document ; l'ontologie de domaine représentant la connaissance dans un domaine spécifique - physique, didactique, épistémologie ; l'ontologie du document se compose d'un modèle de document - organisation et sélection - et d'un modèle d'adaptation. Ce modèle de document et modèle d'adaptation deviennent une ontologie didactique de scénario fondé sur un modèle hiérarchique de tâches et un nouveau modèle d'adaptation fondé sur les catégories d'adaptation nécessitées par le projet de MODALES ; une ontologie utilisateur qui définit les différents stéréotypes - catégories des professeurs et des apprenants et leurs différentes propriétés. Ces connaissances sont fondées sur les connaissances communes et explicites partagées par tous les membres de la Communauté de pratiques. En d'autres termes, les scénarios sont les points clé pour concevoir l'ontologie de document, le schéma de métadonnées et les spécifications - ou scénario générique.

En comparaison de la version en courante de SCARCE, nous avons besoin de modifier/étendre le modèle de document et le modèle d'adaptation et le moteur de composition sémantique - au moins. La nouvelle version du moteur de composition sémantique doit pouvoir contrôler un modèle hiérarchique de tâches et de traiter de nouveau modèle d'adaptation.

5.7 Conclusion

Nous allons maintenant présenter les objectifs et les résultats attendus pour l'étape 2 du projet. A partir de la modélisation (résultats des travaux de l'année I, Tâche I.3), élaboration d'une plate-forme d'EIAH fondée sur un scénario générique capable d'engendrer toutes les situations d'apprentissage identifiées intégrant les paramètres de variabilités suivants : i) la catégorie des apprenants ayant une variabilité intra et inter catégorie ii) les ressources disponibles dans différents domaines - physique, didactique et épistémologie iii) la répartition entre formation en présentiel et à distance iv) la distribution de l'activité entre les différents « participants » (machine, apprenants, tuteurs, enseignants, etc.).

- Tâche II.1 : Formalisation des modèles issus de la Tâche I.3, Année I sous forme d'ontologie dans un format conforme au standard du W3C (OWL, Ontology Web Language).
- Tâche II.2 : Sélection et indexation des ressources d'apprentissage.
- Tâche II.3 : Elaboration de la plate-forme EIAH adaptative, mise en place de nouveaux gabarits de page Web (nouveaux formats de présentation, nouveaux types de ressources) et intégration du modèle de la pédagogie active dans le processus de génération dynamique de cours et intégration des modèles et des ressources.

- Tâche II.4 : Les résultats de la première année seront affinés par la mise en place de l'étude d'une seconde session de formation à destination des mêmes publics selon les procédés décrits dans les Tâches I.1, I.2, I.3 de la 1^{ère} année. La confrontation des résultats et de leur cohérence permettra de faire évoluer et valider la modélisation retenue pour l'élaboration de la plate-forme EIAH adaptative.

Les résultats attendus pour cette étape 2 sont les suivants :

- Définition d'un modèle de tâches hiérarchiques adéquat à représenter les concepts de la théorie anthropologique du savoir de Chevallard, la répartition entre formation en présentiel et à distance et la distribution de l'activité entre les différents « participants » (machine, apprenants, tuteurs, enseignants, etc.), les différentes catégories d'utilisateurs, ainsi que les stratégies d'adaptation permettant de prendre en compte la variabilité des scénarii.
- Définition et implantation des ontologies suivantes : ontologie des utilisateurs, ontologie du domaine, ontologie du schéma de métadonnées et ontologie des scénarii et de l'adaptation. Cette dernière intègre le modèle de tâches hiérarchiques défini ci-dessus, ainsi que la transposition de certains concepts de la théorie de Chevallard.
- Définir un scénario générique capable d'engendrer toutes les situations d'apprentissage identifiées
- Définir l'apport de d'autres théories didactiques à la mise en œuvre des scénarii.
- Evaluer les usages et les apprentissages en jeu dans l'EIA.

5.8 Crédits

Les études menées dans le cadre du projet MODALES doivent beaucoup à l'équipe du CREAD participant au projet dont plus particulièrement à M. Gérard Sensevy, M. Sylvain Laubé, M. Yves Kuster, M. Dominique Forest et Mme Catherine Loisy et à M. Jean-Louis Tetchueng Foping qui effectue sa thèse sur ce projet.

5.9 Bibliographie

- Akrich, M., M. Callob, et al. (1998). "A quoi tient le succès des innovations." Annales des Mines: 4-17.
- Brousseau, G. (1998). Théorie des situations didactiques. Grenoble.
- Chevallard, Y. (1992). "Concepts fondamentaux de la didactique: perspectives apportées par une approche anthropologique." La Pensée sauvage, Grenoble, Recherches en didactique des mathématiques 12(1): 73-112.
- Chevallard, Y. (1999). "L'analyse des pratiques enseignantes en théorie anthropologique du didactique." La Pensée sauvage, Grenoble, Recherches en didactique des mathématiques 19(2): 221-226.
- Cole, M. and Y. Engeström (1993). A cultural-historical approach to distributed cognition. Distributed cognition. Psychological and educational considerations. G. Salomon, Cambridge University Press: 1-46.

- Derycke, A., V. Chevrin, et al. (2005). Intermédiations Multicanales et Multimodales pour l'E-formation : l'architecture du projet Ubi-Learn. EIAH 2005, Montpellier.
- Derycke, A., F. Hoogstoel, et al. (2006). Infrastructure des environnements informatiques pour l'apprentissage distribué. à paraître. J. Labat and M. Grandbastien, Hermès Sciences Publications.
- Derycke, A., Y. Peter, et al. (2005). "Conception d'une Plate-forme pour la recherche en EIAH." STICEF 12.
- Duursma, C., O. Olsson, et al. (1993). Task model definition and task analysis process. Deliverable ESPRIT Project P5248 KADS-II/M5/VUB/RR/004/2.0. T. R. CK-VUB-04, Free University Brussels.
- Garlatti, S. and S. Iksal (2004). A Flexible Composition Engine for Adaptive Web Sites. Adaptive Hypermedia and Adaptive Web-Based Systems, AH 2004. P. D. Bra and W. Nejdl, Springer Verlag. LNCS 3137: 115-125.
- Garlatti, S., S. Iksal, et al. (2004). SCARCE: an Adaptive Hypermedia Environment Based on Virtual Documents and Semantic Web. Adaptable and Adaptive Hypermedia Systems. S. Y. Chen and G. D. Magoulas., Idea Group Inc.: 206-224.
- Garlatti, S., J.-L. Tetchueng, et al. (2006). The Co-Design of Scenarios for a Didactic-based E-learning System viewed as an Adaptive Virtual Document. Satellite Workshop E-Learning, 2nd IEEE International Conference on Information & Communication Technologies : from Theory to Applications, Damascus, Syria.
- Iksal, S. and S. Garlatti (2004). Adaptive Web Information Systems: Architecture and Methodology for Reusing Content. AH 2004 workshops, EAW'04: Engineering the Adaptive Web, Eindhoven, Technische Universiteit Eindhoven.
- Laubé, S., S. Garlatti, et al. (2006). Scénarios intégrant les TICE : les méthodologies et les cadres théoriques à l'œuvre dans la recherche MODALES. 8ème Biennale de l'Education, Colloque "Scénariser l'enseignement et l'apprentissage : une nouvelle compétence pour le praticien ?" INRP - ERTé E-Praxis, Paris.
- Laubé, S., S. Garlatti, et al. (2006). Modelling Teachers Activities to Design Scenarios for a Didactic-based E-learning System viewed as an Adaptive Virtual Document. European Conference on Education Research, Genova.
- Laubé, S., S. Garlatti, et al. (2005). Formations des maîtres en sciences : de la modélisation des situations didactiques comme fondement de la conception d'un EIAH adaptatif. EIAH 2005, Montpellier.
- Laubé, S., S. Garlatti, et al. (2005). Formations des maîtres en sciences: De la modélisation des situations didactiques comme fondement de la conception d'un Environnement Informatique pour l'Apprentissage Humain (EIAH) adaptatif. 4èmes rencontres de l'ARDIST, INRP.
- Laubé, S., Y. Kuster, et al. (2006). MODALES : Conception de scénario pour générer des EIAH adaptatifs pour la formation des maîtres en sciences expérimentales. Colloque " Scénarios, Formation d'enseignants : quels scénarios ? quelles évaluations ?" IUFM de Versailles.
- Perrenoud, P. (1998). "La transposition didactique à partir de pratiques : des savoirs aux compétences." Revue des sciences de l'éducation, Montréal XXIV(3): 487-514.
- Sensevy, G., A. Mercier, et al. (2000). "Vers un modèle de l'action didactique du professeur. A propos de la course à 20." La Pensée sauvage, Grenoble, Recherches en didactique des mathématiques 20(3).
- Tchounikine, P. (2002). "Pour une ingénierie des Environnements Informatiques pour l'Apprentissage Humain." Revue I3 2(1): 59-95.

- Tchounikine, P. and Al. (2004). Platon-1: quelques dimensions pour l'analyse des travaux de recherche en conception d'EIAH. Rapport de l'action spécifique "Fondements théoriques et méthodologiques de la conception des EIAH", Département STIC, CNRS.
- Trichet, F. (1998). DSTM: un environnement de modélisation et d'opérationnalisation de la démarche de résolution de problèmes d'un Système à Base de Connaissances. Nantes, Université de Nantes.
- Tricot, A. and F. Plegat-Soutjis (2003). "Pour une approche ergonomique de la conception d'un dispositif de formation à distance utilisant les TIC." Revue STICEF(10).
- Wenger, E. (1998). Communities of Practice - Learning, Meaning and Identity. Cambridge, Cambridge University Press.
- Wielinga, B., W. V. d. Velde, et al. (1992). The KADS Knowledge Modelling Approach. Proceedings of the 2nd Japanese Knowledge Acquisition for Knowledge-Based Systems Workshop. R. Mizoguchi and H. Motoda. Hitachi, Advanced Research Laboratory, Hatoyama, Saitama, Japan: 23-42.
- Willamowski, J. (1992). Modélisation de tâches pour la résolution de problèmes en coopération système-utilisateur. Grenoble, Université Joseph Fourier.

Chapitre 6. Conclusion et perspectives

Après une phase initiale de recherche centrée sur la représentation de connaissances et le raisonnement, la thématique a été étendue vers la conception de systèmes d'information sémantiques et adaptatifs. L'objet principal de cette recherche est de définir une classe de méthodologies adaptées à la conception de ces systèmes d'information. Ces derniers doivent assurer une distribution du rôle de l'interprétation et de l'action entre l'homme et la machine dans une communauté de pratiques particulières. L'activité de recherche s'est organisée autour de quatre étapes significatives qui se sont constituées autour de projets concrets avec une approche pluridisciplinaire définissant un contexte d'interprétation et d'action dans une communauté de pratiques particulière. Dans chaque étape, le ou les projets concernés partagent des problématiques, des méthodes et des technologies communes qui retracent les évolutions significatives de l'activité de recherche.

Ce chapitre est organisé en deux parties principales : l'une présente de manière synthétique les quatre étapes de recherche en décrivant les différentes problématiques de recherche, leurs caractéristiques et les contributions correspondantes et l'autre présente les perspectives organisées en cinq catégories : approfondissement des problématiques actuelles, technologiques et logiciels, valorisation, organisation de la recherche et projets futurs illustrant ces perspectives.

Le tableau 3, de la page suivante, présente une vue synthétique des différentes problématiques de recherche et des contributions correspondantes. Le tableau 4 présente les caractéristiques principales des systèmes d'information pour chaque étape

Tableau 3 : Problématiques de recherche des différentes étapes et leurs contributions

Etapes	Problématiques de l'étape et des projets		Contributions
1	<ul style="list-style-type: none"> Analyse de tâches expertes : diagnostic médical Mise en correspondance des connaissances médicales et des connaissances en situation Ressources locales, maîtrisées par les concepteurs et/ou utilisateurs 	<ul style="list-style-type: none"> ATLAS : Aide au diagnostic : accès à des informations pertinentes en fonction du contexte pour différentes catégories de médecins Représentation des connaissances et mécanismes d'inférence associés ATLAS-TUTOR : EIAH d'aide au diagnostic et détection d'erreurs de diagnostic et remédiation pour différents niveaux d'expertise représentation des connaissances pertinentes pour l'aide au diagnostic 	<ul style="list-style-type: none"> Réalisation d'un hypermédia contextuel : navigation contextuelle pour la recherche d'information avec un contexte défini par un domaine médical et/ou une tâche Proposition d'une base de connaissances intensionnelle fondée sur des points de vue munis d'un mécanisme d'héritage multiple avec conflits
2	<ul style="list-style-type: none"> Hypermédia adaptatif Ressources locales, maîtrisées par les concepteurs et/ou utilisateurs 	<ul style="list-style-type: none"> SWAN : Conception d'un système d'information sémantique et adaptatif Acquisition et organisation des connaissances pour l'adaptation et les activités utilisateur 	<ul style="list-style-type: none"> Apports de l'intégration d'un tuteur intelligent et d'un ATLAS ayant des connaissances anatomiques et spatiales Nouvelles actions d'apprentissage pour la remédiation fondées sur ces connaissances spatiales et anatomiques
3	<ul style="list-style-type: none"> Document virtuel adaptatif (DVA) Réutilisation du DVA Grande masse de ressources hétérogènes et géographiquement distribuées Production et description des ressources faites par une communauté géographiquement distribuées 	<ul style="list-style-type: none"> CANDLE – ICCARS – ERGOWEB - KMP : Système d'information sémantique et adaptatif Acquisition et organisation des connaissances pour l'adaptation et la gestion de la cohérence Réutilisation, partage et échange des ressources distribuées Production de résultats cohérents 	<ul style="list-style-type: none"> Proposition d'une méthodologie de conception d'un système d'information adaptatif fondée quatre modèles sémantiques : <ul style="list-style-type: none"> Un modèle de tâches hiérarchiques pour l'organisation Un filtrage spatial part tâche de recherche d'information, Un modèle de domaine pour l'indexation, Un modèle de contexte et un modèle utilisateur pour l'adaptation. Proposition d'une méthodologie de conception des SIAS : <ul style="list-style-type: none"> Réalisation d'un moteur de composition (SCARCE), flexible et réutilisable fondé sur le web sémantique et muni de spécifications déclaratives décliné et adaptée à différents contextes Spécifications de l'organisation, de la sélection et de l'adaptation fondée sur des ontologies. Gestion de la cohérence : connaissances explicites de la communauté de pratiques par des ontologies : domaine, document (structure narrative et adaptation), utilisateur et métadonnées Principes de SCARCE réutilisés pour SCARCE^{WS} et le projet Knossos Transfert technologique vers la société Atlantide
4	<ul style="list-style-type: none"> Problématiques de l'étape 3 Evaluation de l'EIAH Construction d'une communauté de pratiques 	<ul style="list-style-type: none"> MODALES : EIAH adaptatif Problématiques de l'étape 3 Traçabilité des modèles et des théories 	<ul style="list-style-type: none"> Proposition d'une méthodologie d'acquisition de scénarios d'apprentissage : <ul style="list-style-type: none"> Savoir-faire des formateurs, Théorie didactique anthropologique de Chevallard et modèle de tâches hiérarchiques Réalisation d'une extension de SCARCE

Tableaux 4 : Caractéristiques principales des systèmes d'information pour chaque étape

Etapes / Projets	Indexation / réutilisation	Sélection : recherche d'information	Organisation et activités utilisateur correspondante		Adaptation	Technologies
Etape 1 : ATLAS, ATLAS TUTOR	<ul style="list-style-type: none"> • Points de vue : Modèles de domaines • Réutilisation des ressources : locale 	<ul style="list-style-type: none"> • Filtrage des sous-domaines dans des Tâches • Requêtes par navigation dans base données et base de connaissances 	<ul style="list-style-type: none"> • Modèles de domaines 	<ul style="list-style-type: none"> • Diagnostic médical 	<ul style="list-style-type: none"> • Non 	<ul style="list-style-type: none"> • Lelisp, Yafool, Y3
Etape 2 : SWAN	<ul style="list-style-type: none"> • Modèles de domaine • Réutilisation des ressources : locale 	<ul style="list-style-type: none"> • Filtrage spatial en fonction de la tâche • Requête à la base de connaissances 	<ul style="list-style-type: none"> • Modèles hiérarchiques de tâches 	<ul style="list-style-type: none"> • Activités partiellement décrites par le modèle tâches hiérarchiques 	<ul style="list-style-type: none"> • Adaptation gérée par <ul style="list-style-type: none"> • Modèle utilisateur • Modèle de contexte 	<ul style="list-style-type: none"> • Lispworks, • Loom, • Java • WebObjects
Etape 3 : ICCARS, CANDLE, ERGOB WEB, KMP	<ul style="list-style-type: none"> • Schéma de Métadonnées • Réutilisation des ressources sur internet 	<ul style="list-style-type: none"> • Requêtes à un système de gestion de connaissances 	<ul style="list-style-type: none"> • Structure narrative fondée sur le savoir-faire des auteurs (journalistes et enseignants) 	<ul style="list-style-type: none"> • Lecture et compréhension d'un dossier thématique ou d'un cours en ligne • Production et description de ressources 	<ul style="list-style-type: none"> • Modèle d'adaptation déclaratif • Principe unique d'adaptation 	<ul style="list-style-type: none"> • Apache, Tomcat, Java, XML, XSLT • Web sémantique : RDF/RDFS, OWL (F-logic), • Ontobroker
Etape 4 : MODALES	<ul style="list-style-type: none"> • Schéma de Métadonnées • Réutilisation des ressources sur internet 	<ul style="list-style-type: none"> • Requêtes à système de gestion de connaissances 	<ul style="list-style-type: none"> • Modèle de tâches hiérarchiques munies de méthodes, • Catégories de tâches provenant de la théorie didactique de Chevallard et du savoir-faire des enseignants 	<ul style="list-style-type: none"> • Apprentissage par résolution de problèmes : Réalisation d'un cours sur l'existence de l'air 	<ul style="list-style-type: none"> • Modèle d'adaptation déclaratif, • Principes d'adaptation : <ul style="list-style-type: none"> • Tâches / Méthodes • Ressources 	<ul style="list-style-type: none"> • Apache, Tomcat, Java, XML, XSLT • Web sémantique : RDF/RDFS, OWL (F-logic), • Ontobroker

Comme on peut le voir dans le tableau 3, les technologies et logiciels utilisés lors de ces différentes étapes ont très largement évolué. Certains logiciels et technologies utilisés pour les deux premières étapes ne sont plus d'actualité, soit parce qu'ils ne sont plus maintenus, soit parce qu'ils ne permettent plus d'assurer les contraintes nécessaires aux systèmes d'information actuels. L'une des principales causes est la nécessité d'utiliser des standards pour assurer l'interopérabilité et la réutilisation. Cette contrainte est aussi un avantage puisqu'elle permet d'envisager de manière plus pérenne la création et la maintenance de plateformes logicielles réutilisables et extensibles pour l'avenir, notamment pour SCARCE. En effet, les développements précédents SCARCE n'ont pu être maintenus malgré notre volonté d'assurer leur réutilisation. Les investissements humains nécessaires au développement de nouveaux environnements sont très coûteux.

Ces différentes étapes de recherche ont permis d'apporter des contributions à différentes communautés scientifiques sans épuiser les problématiques de recherche présentées dans ce rapport. En effet, il reste encore de nombreux aspects à approfondir. Cet approfondissement ne peut se faire sans prendre en compte les dimensions suivantes :

- Les technologies et plateformes logicielles.
- La valorisation.
- L'organisation de la recherche.

Dans premier temps, nous allons présenter l'approfondissement des problématiques de recherche. Ce dernier va se poursuivre selon six axes majeurs. Ces six axes sont déjà instanciés aux travers de projets qui ont commencé ou qui vont commencer en 2007. Ces six axes majeurs ne sont pas indépendants comme nous le verrons plus tard. Ils forment un tout cohérent. Ces axes principaux de développement sont les suivants :

1. Extensibilité de SCARCE :

On a vu précédemment que l'environnement SCARCE possède des limitations. En effet ces propriétés fondamentales se retrouvent dans le modèle de document et d'adaptation, ainsi que dans les règles de composition du moteur. Pour le projet KMP, il est nécessaire de changer le modèle de document et d'adaptation, donc les règles de composition du moteur. Il serait intéressant d'analyser et d'étudier les modèles de document et d'adaptation assurant la meilleure extensibilité possible dans SCARCE, ainsi qu'une étude plus fine de cette notion de propriétés fondamentales d'un moteur de composition.

2. Rapport coût / gains de la description du sens, de la distribution de l'interprétation et de l'action entre l'homme et la machine :

Indépendamment de la conception des schémas de métadonnées et des ontologies correspondantes, il y a un problème de saisie des métadonnées en général peut-on automatiser ou semi automatiser cette saisie ? en cas de saisie automatique des métadonnées quelle est la limite en termes d'automatisation des tâches machines ? En effet, dans certains secteurs d'application, il n'est pas pensable de saisir des métadonnées manuellement ou alors une très faible partie d'entre elles, comme par exemple dans le

domaine des programmes audiovisuels appliquée à de nouveaux modes d'accès délinéarisé et personnalisé qui seront distribués au travers d'internet et sur différents supports multimédia.

3. Web services :

La conception des systèmes d'information du futur un système d'information peut être vu comme un moteur de composition adaptatif qui sélectionne, organise et assemble dynamiquement des ressources et des web services multimédia pertinents. Il est généré dynamiquement à partir de modèles sémantiques – ontologies. On voit bien ici que dans cette définition, il n'est pas seulement question de composer des ressources statiques mais aussi des ressources que nous appellerons dynamiques qui sont les web services. En effet, il n'est plus uniquement nécessaire d'assurer la réutilisation des ressources statiques, mais aussi la réutilisation de web services pour assurer une meilleure flexibilité et extensibilité. Le système d'information aura toujours une partie interaction homme/machine qui prendra en compte les besoins utilisateur comme actuellement pour la composition de ressources statiques, mais le contenu affiché à l'utilisateur pourra très bien être le résultat fourni par un web services distant ou par la composition de web services distribués.

4. Stratégies d'adaptation :

Actuellement, SCARCE possède des stratégies d'adaptation limitées. En effet, il ne met en œuvre que des méthodes de navigation adaptative provenant de la communauté hypermédia adaptatifs (navigation et présentation adaptative). Cette expression de l'adaptation est d'une certaine manière de bas niveau. Un système d'information adaptatif doit permettre d'adapter les ressources, l'organisation et la présentation. En effet, elles sont proches de la mise en œuvre d'un système d'information et pas des connaissances et savoir-faire des utilisateurs. Il faut être capable de déclarer différentes stratégies d'adaptation d'un niveau d'abstraction supérieur, d'en faire des catégories munies de leurs propres stratégies afin qu'elles soient proches du niveau de compréhension des utilisateurs et de leur savoir-faire et ainsi favoriser leur utilisation, maîtrise, diffusion et partage dans la communauté de pratiques visée. L'adaptation dynamique de SCARCE est également limitée. L'adaptation est dite dynamique lorsque le modèle utilisateur est modifié à partir des observations de son comportement. Actuellement, l'adaptation dynamique n'est utilisée que pour régénérer le document hypermédia lors d'une nouvelle session. En effet, le document engendré pour un utilisateur donné n'a pour durée de vie que celle de la session ; il n'est pas rémanent. Dans tout système adaptatif et plus particulièrement pour un EIAH, il est nécessaire d'avoir un observateur des utilisateurs qui mémorisent certains paramètres pertinents pour ensuite mettre à jour le modèle utilisateur avant d'engendrer la prochaine étape du document - cours. L'adaptation peut être vue comme un quadruplet : ressources, modèle utilisateur, observateur et règles d'adaptation. Ce type de quadruplet a été utilisé pour une formalisation logique de l'adaptation afin de comparer différents EIAH fondés sur des hypermédiats adaptatifs. En effet, le rôle de l'adaptation est de fournir au bon moment les ressources pertinentes à un apprenant - en termes d'apprentissage. Les règles d'adaptation comparent une description des ressources – métadonnées – au modèle utilisateur pour décider de leurs pertinences. En cela, l'adaptation peut être vue comme un raffinement de la sélection ou recherche d'information. Avant d'appliquer les règles d'adaptation, l'observateur a enregistré

certaines propriétés du comportement de l'apprenant et s'en sert pour modifier son modèle. Un modèle d'interaction déclaratif est donc un modèle qui permet l'explicitation de la boucle : Interaction (de l'apprenant), Observation (par le système et mise à jour du modèle utilisateur) et Adaptation (application des règles) en relation avec le modèle utilisateur et le schéma de métadonnées. Un tel modèle doit permettre de déclarer des catégories d'adaptation auxquelles sont associées des observations et des règles d'adaptation qui sont bien évidemment liées aux métadonnées et modèle utilisateur.

5. Adaptation et contexte :

De nos jours, la présence quasi omniprésente des réseaux de télécommunications nous permet d'envisager de nouvelles organisations des services et des usages. Il s'agit de proposer de nouveaux services disponibles quelque soit l'endroit, le réseau, le terminal, etc. Pour cela, il est nécessaire de prendre en compte le contexte social, technologique, organisationnel, réseaux, terminaux, etc. qui est évolutif et dynamique. Il faut donc être capable de s'adapter dynamiquement aux besoins utilisateur et aux différents contextes.

6. Composition de systèmes d'information adaptatifs :

Actuellement de nombreuses équipes travaillent sur la conception de systèmes d'information adaptatifs qui autorisent nécessairement la réutilisation de ressources distribuées - statiques ou dynamiques comme les web services. Il est bien évident qu'il va être nécessaire d'assurer la réutilisation de systèmes d'information adaptatifs qui devront être interopérables pour pouvoir être composés. On voit bien que cette nouvelle fonctionnalité a une implication directe sur la définition des systèmes d'information du futur, mais aussi que cette fonctionnalité est très liée à la réutilisation de web services, en effet comment traiter ceci sans passer par des web services, même si ces derniers ne sont pas suffisants pour résoudre le problème. Mais la réutilisation des web services adaptatifs reste encore un verrou technologique. De nombreux débats ont eu lieu dans la communauté Hypermédia Adaptatif sur ce sujet, mais sans solution probante pour l'instant. On peut ré-exprimer le problème de la manière suivante : comment doit-on décomposer un système d'information en web services – problème de granularité - pour assurer la réutilisation de services adaptatifs ? En effet, si l'on désire réutiliser un web service adaptatif, il est nécessaire de « connaître » et/ou « disposer » de l'information et de connaissances sur les différents éléments du quadruplet « ressources, modèle utilisateur, observateur, règles d'adaptation » pour concevoir un environnement distribué, fondé sur des web services. Il s'agira notamment de trouver le bon niveau de granularité des web services pour assurer cette réutilisation.

Tous ces axes de développement sont présents dans des projets à venir qui sont cités à la fin de ce chapitre. Nous allons maintenant aborder les technologies et les plateformes logicielles.

L'une des contributions de cette activité de recherche est un moteur de composition adaptatif, flexible et réutilisable. Grâce aux propriétés de flexibilité et de réutilisabilités de ce moteur de composition, il est tout à fait envisageable de concevoir de très nombreuses classes de systèmes d'information adaptatifs et sémantiques à l'aide ce moteur. En d'autres termes, de très nombreuses applications accessibles sur internet pourrait être conçus à l'aide ce moteur. Deux objectifs principaux sont liés à ce moteur de composition : l'un est d'assurer la pérennité de cet environnement afin d'assurer sa réutilisation dans différents contextes, mais

aussi d'en faire une plateforme logicielle libre et l'autre est d'effectuer un transfert technologique vers l'industrie soit du moteur de composition et de la méthodologie associée, soit uniquement de la méthodologie lorsque les technologies sous jacentes ne sont pas entièrement transposables actuellement.

Du point de vue de la valorisation, il est nécessaire d'améliorer la visibilité de ces activités tant d'un point de vue académique par la publication d'articles de revue, que du point de vue des communautés scientifiques par une meilleure présence et développement de ces problématiques dans des communautés comme celles de RFIA (Reconnaissance des Formes et Intelligence Artificielle), DVP (Document Virtuels Personnalisables), EIAH, Adaptive Hypermedia, Web Intelligence, Semantic Web, Pervasive or Ubiquitous Computing and learning, E-learning, etc.

Toutes ces perspectives ne peuvent se faire sans penser à l'organisation de la recherche. AU sein du GET (groupement des écoles des télécommunications), la recherche est organisée en projets structurants. Ce sont des projets ayant une thématique de recherche bien identifiée, qui sont notamment constitués de personnels permanents et non permanents associés (thésards et post-doctorants) et des projets particuliers financés par l'Europe, l'état, la région et des industriels. Actuellement responsable du projet structurant SCRIPTureS, l'organisation de la recherche à court et moyen terme va donc passer par celle de ce projet. Ma mutation au sein du département informatique va permettre de développer de nouvelles collaborations au sein de ce département et d'accroître le potentiel de ce projet, notamment par certaines fusions entre projets. Au sein du GET, il est également important de développer des collaborations internes tant du point de vue des problématiques de recherche que du point des plateformes logicielles afin d'accroître la visibilité du GET dans ces domaines. Ces collaborations ont déjà commencé avec le projet SIMBAD (Responsable Bruno Defude) et vont se poursuivre et s'intensifier – projets SIAS et P-Learnnet.

Il est également important de développer la communauté française autour du « document virtuel ». Dans cette communauté, des conférences et des réunions de travail pour son organisation ont déjà été faites. Mais, il est nécessaire de développer sa visibilité à l'international – c'est un des enjeux mentionnés lors de ces réunions de travail.

Une collaboration sur nos problématiques de recherche a été initialisée avec le Vietnam. Elle s'est notamment concrétisée par la présence d'un enseignant de l'Université de Ho Chi Minh Ville en thèse dans notre laboratoire. Cette collaboration sera donc poursuivie et intensifiée dans l'avenir.

Nous allons maintenant présenter les projets qui instancient les six axes majeurs d'approfondissement des problématiques de recherche. La présentation de ces projets est extraite des propositions de soumission de ceux-ci et sont donc le résultat d'un travail collectif.

Semim@ge

Ce projet va permettre d'étudier le rapport coût / gains de la description du sens et d'étendre SCARCE par un ou des modèles de document et leurs règles de composition dans le domaine audiovisuel. Le but du projet SemIm@ge est l'exploration et le prototypage

d'applications d'accès et de consommation délinéarisée, personnalisée de contenus audiovisuels professionnels. Pour cela, il vise à développer une solution de re-détermination (repurposing) et de ré-éditorialisation de programmes audiovisuels professionnels sur la base exclusive d'une chaîne de traitement la plus automatique possible et générique. Ce système d'information proposera une publication adaptée à différents vecteurs de diffusion et des services associés en passant par l'indexation par génération automatique de métadonnées sémantiques et structurelles.

Ce projet est principalement motivé par les besoins issus des évolutions fondamentales que vit actuellement le secteur de la diffusion des contenus audiovisuels :

- La multiplication des canaux de distribution, avec notamment distribution sur des réseaux possédant la propriété d'interactivité et permettant de ce fait le développement croissant d'une consommation audiovisuelle de plus en plus personnelle et à la carte, donc délinéarisée, associée à un accroissement de l'usage de l'Internet au détriment du temps passé à une consommation passive de la TV.
- Le lancement de moteurs de recherche Web dédiés à la vidéo TV par plusieurs acteurs majeurs des moteurs de recherche tels que Google ou Yahoo.
- La baisse prévisible des revenus publicitaires de la diffusion TV linéaire due notamment à une fragmentation croissante du marché.

Il s'agit donc d'offrir des contenus TV par une distribution adaptée et diversifiée en termes de services offerts sur les différents canaux possibles. On peut prévoir un intérêt croissant pour des solutions de « repurposing » à coût raisonnable c'est-à-dire fondées essentiellement sur des traitements automatiques et une utilisation massive de métadonnées sémantiques. Pour atteindre cet objectif, des algorithmes de reconnaissance automatiques d'usage relativement universel guidés par un modèle de connaissance spécifique de la famille de contenus à analyser seront utilisés. Ces connaissances, notamment décrites par des ontologies, ont essentiellement trois rôles :

- Paramétrer l'analyse automatique (exemple : appliquer le bon modèle de langage d'un moteur de transcription vocale).
- Décrire les connaissances formelles permettant une forme plus ou moins élaborée d'interprétation du contenu (comme par exemple la structure du contenu) mais aussi les raisonnements permettant de prendre une décision conduisant à une description (ou indexation) structurée des contenus - métadonnées.
- Définir les règles de composition et d'adaptation aux utilisateurs qui permettront la réutilisation et l'organisation de contenus pour une « ré-éditorialisation ». Ces règles de composition seront fondées sur les connaissances explicites - savoir-faire - des communautés de pratiques du domaine.

Le caractère novateur du projet réside d'une part dans un couplage étroit des fonctions de gestion, moteur de recherche et de « repurposing » fondé d'une part sur la production automatisée de métadonnées sémantiques et d'autre part sur la complétude des sujets traités et maîtrisés par le consortium sur le plan des algorithmes de reconnaissance audio, vocale et

visuelle, de traitement des connaissances et de composants système pour arriver à ce résultat et envisager une expérimentation en vraie grandeur sur un volume de contenus conséquent.

Ce projet va notamment étudier les points suivants :

- La publication chapitrée de programmes TV (informations, sport, variétés) pour l'Internet et les mobiles. Actuellement, ce type de publication est réalisée quasiment entièrement par une « ré-éditorialisation » humaine avec des environnements souvent spécifiques.
- La publication et la diffusion personnalisée d'informations, en fonction de thèmes choisis par l'utilisateur, ceci à travers un ensemble de chaînes disponibles.
- La production automatique de résumés de programmes d'information ou de sports (temps forts, points marqués);
- L'édition de best-of de personnalités connues.

Ce projet a pour objet de proposer ces fonctionnalités de (re-) publication sur la base de traitements les plus génériques possible et d'un système entièrement paramétrable par des connaissances caractérisant d'une part la structure du programme à analyser et d'autre part le type de publication souhaitée. Pour cela, elle repose sur les points essentiels suivants :

- Utilisation de technologies de reconnaissance automatique dans l'image et l'audio permettant la production de métadonnées sémantiques. Pour cela les technologies suivantes seront réutilisées et adaptées :
 - de classification des sons en parole – musique ou divers.
 - de transcription vocale.
 - d'identification d'œuvres musicales.
 - de macro-segmentation multimodale de flux TV basées sur le regroupement automatique d'éléments visuels et sonores similaires.
 - de reconnaissance de textes incrustés dans les vidéos.
 - de reconnaissance de visages dans les vidéos.
 - de traitement automatique du langage naturel, notamment dans un but de production de vocabulaires thématiques.
- Production d'une description structurée par catégorie d'émission formalisée par une ontologie décrivant la famille de programmes en question et les raisonnements correspondants. Ces connaissances permettront la mise en œuvre d'une analyse multimodale croisant les caractéristiques sonores et visuelles.
- Republication qui réutilise les contenus décrites par des métadonnées, les filtre – sélection des contenus les plus pertinents, puis les organiser en fonction des

connaissances de la communauté de pratiques et des besoins utilisateurs - adaptation. Ces connaissances permettent d'expliciter les « règles de composition » des contenus sélectionnés afin d'assurer cette republication.

- Intégration de l'ensemble de ces éléments au sein d'une chaîne complète automatisée permettant un fonctionnement opérationnel en continu sur plusieurs semaines pour une évaluation en vraie grandeur. Pour cela, un certain nombre de fonctions préexistantes chez certains partenaires comme l'acquisition, l'enregistrement et le transcodage automatisés seront également utilisées.

Les contenus considérés pourront provenir soit de l'enregistrement de flux TV, soit par ingestion de programmes archivés sous forme analogique ou numérique. Dans le cas des contenus TV, des métadonnées préexistantes telles que les données EIT (Event Information Table du standard DBV-SI) véhiculant des informations de guide de programme pourront également être exploitées.

SIAS

Ce projet aborde les points suivants : les stratégies d'adaptation, les Web services, la composition de systèmes d'information adaptatifs et l'extensibilité de SCARCE. L'évolution actuelle des systèmes d'information (SI) va vers plus de personnalisation et de réutilisation dans un environnement informatique marqué par la répartition et les technologies du web (dont les web services). La notion de composition de ressources (documents, présentations, composants, web services, ...) devient centrale. Cette composition permet la construction de systèmes complexes par assemblage de composants élémentaires. Cette dernière devient de plus en plus dynamique et adaptative, c'est-à-dire dépende du contexte dans lequel le système est déployé. La notion de contexte est très large et peut inclure l'environnement matériel et logiciel (y compris le réseau), mais aussi l'utilisateur (sous différentes facettes). Ceci ne pourra se faire sans une réelle modélisation tant du contexte, que des composants et de la composition elle-même. A notre sens, cette modélisation doit décrire non seulement les aspects syntaxiques, mais aussi les aspects sémantiques. Cette sémantique est généralement décrite par des métadonnées et les ontologies correspondantes. Il devient donc possible d'automatiser et/ou de semi automatiser la création des SI. L'idée de ce projet est donc de chercher à mieux analyser la notion de composition pour être capable de définir ensuite des moteurs de composition beaucoup plus réutilisables, extensibles et flexibles.

Actuellement, les communautés Adaptive Hypermédia, Web services (W3C, Web sémantique et Web Intelligence), documents numériques (ACM Symposium on Document Engineering, RTP 30 Documents), documents virtuels, E-learning (RTP 39, réseaux d'excellences Kaléidoscope et Prolearn) abordent une problématique commune sous des angles différents : le besoin de réutiliser des ressources (documents, présentations, composants, web services, ...) indexées par des métadonnées sémantiques et les ontologies correspondantes pour la conception des nouveaux systèmes d'informations fondés sur le web sémantique et les standards du W3C. La notion même de moteur de composition devient dès lors centrale et permet d'aborder le problème de manière plus conceptuelle. On est donc amené à se poser des questions fondamentales telles que : définition de la notion de composition, principes fondamentaux, liens entre principes et domaines d'application, conditions de réutilisation du moteur de composition, maintenance d'une application à l'aide

de ce type d'approche, etc. Une analyse approfondie de la notion de composition aura donc un impact non négligeable sur ces différentes communautés. Actuellement, chaque communauté est concentrée sur sa problématique propre et ses mécanismes particuliers de composition sans aborder nécessairement le problème de fond du « moteur de composition ». L'innovation de ce projet réside donc dans l'étude même de la notion de composition, de la réutilisation et de la flexibilité des moteurs de composition et de ses impacts sur la conception des systèmes d'information de demain.

P-LEARNET

Ce projet aborde les points suivants : les stratégies d'adaptation et de gestion du contexte, les Web services, la composition de systèmes d'information adaptatifs et l'extensibilité de SCARCE.

Avec les développements rapides des réseaux de télécommunications offrant la convergence (par des protocoles d'IP) et du « roaming » universel pour les communications sans fil, de GSM / UMTS vers le « WiFi », et ceux des mobiles communiquant (Smartphone, PDAs, etc.), il s'avère que de nouveaux modèles d'organisation et d'infrastructure pour des services et des usages soient envisageables. Ils doivent bien évidemment être soigneusement conçus et évalués des usages pour des applications particulières. Le but principal du projet de p-LearNet est d'explorer le potentiel des communications omniprésentes pour l'apprentissage humain individuel ou collectif, quelque soit l'endroit, la période et les contextes organisationnels et technologiques. Pour cela, il va être nécessaire de prendre en compte le contexte qui va être évolutif ou dynamique en termes de réseaux, de terminaux, etc., mais aussi d'adaptation multi-niveaux. Il semble que l'apprentissage humain, considéré comme des échanges et des constructions de connaissances formelles ou informelles à l'aide de systèmes d'apprentissage mis en valeur par les nouvelles technologies soit un enjeu majeur pour le développement non seulement du niveau d'éducation d'une population, mais aussi pour la compétitivité des organisations. L'originalité de ce projet est d'entreprendre une étude exploratoire sur la mobilité en analysant et combinant les services et les usages dans un cadre d'intelligence ambiante. Le projet va porter sur des utilisations et des besoins réels des entreprises et des organismes du consortium. En effet, il est bien clair que des problèmes génériques tels que l'adaptation multi niveaux et l'adaptation au contexte ne puissent pas être résolus indépendamment d'une application. Les deux enjeux principaux de p-LearNet sont :

- L'évaluation des usages dans un contexte d'informatique omniprésente, ce qui n'est pas une tâche facile, notamment pour collecter les traces expérimentales
- La complexité et la cohérence des adaptations multi niveaux et au contexte de bout en bout.

P-LearNet prendra en compte l'ubiquité des systèmes d'apprentissage, la mobilité des étudiants, les nouveaux besoins nécessaire à ce nouveau mode d'apprentissage, le potentiel des multimédia, des interfaces utilisateur multicanaux et multimodales sensibles à la personnalité de l'utilisateur/de l'apprenant. Le terme « multicanaux » est ici pris dans le sens d'une combinaison riche d'un terminal particulier avec différentes catégories de réseaux (qualité des services, topologie, protocoles, coûts, etc.) incluant des communications vocales et visuelles de personne à personne. Les systèmes d'apprentissage omniprésents impliquent

également la capacité à maîtriser les processus de découverte, de composition, d'adaptation des services contextuels et des interactions avec les utilisateurs. Il est aussi nécessaire de surmonter le problème de la complexité liée aux nombreuses combinaisons possibles de ces différents éléments. P-LEARNET étudiera des problèmes d'adaptation au contexte et multi niveaux de bout en bout. La question d'équilibrer des adaptations locales et globales est centrale.

L'adaptation et la gestion du contexte peut être vu comme quadruplet : le modèle des métadonnées, un modèle d'utilisateur, un modèle de contexte et modèle d'adaptation. Ainsi, les règles d'adaptation doivent « comparer » / « évaluer » les métadonnées avec le contexte et le modèle d'utilisateur pour choisir les ressources pertinentes. Ces modèles devraient être décrits au niveau sémantique pour assurer l'interopérabilité, la flexibilité et la réutilisabilité. Les enjeux principaux sont : Comment contrôler l'adaptation dynamique au contexte et à différents niveaux de bout en bout en préservant les tâches d'apprentissage ou les intentions didactiques/pédagogiques malgré les changements ? Comment contrôler la réutilisation des services web adaptatifs ? Comment obtient-on les informations sur le contexte et comment contrôler les changements et leurs conséquences pour les utilisateurs, en continu ? La gestion du contexte est un processus dynamique avec des dépendances historiques. Elles sont nécessaires pour contrôler les changements et les transitions en continu. Comment contrôler l'adaptation, les modèles dynamiques du contexte et les apprenants d'une manière déclarative ?

REACTIONS

Le but principal du projet de REACTIONS (REuse And Composition of adapTive InformatiON Systems) est d'étudier le problème de la réutilisation et de la composition des systèmes adaptatifs et de créer un groupe européen capable de résoudre ce problème. Pour atteindre cet objectif, le projet se concentrera sur un secteur spécifique : les systèmes de E-Learning.

De nos jours, beaucoup d'entreprises emploient des systèmes d'information en ligne pour leurs clients et leurs employés. Internet fournit à ces clients et ces employés une quantité énorme d'information, qui est fournie partout et à tout moment. Cependant, tous les utilisateurs n'ont pas les mêmes objectifs, connaissances, centres d'intérêts et besoins. Par conséquent, il est nécessaire d'adapter l'information, sa présentation et son organisation aux besoins utilisateur et au contexte courant dans le cadre d'une informatique omniprésente. Le besoin d'adaptation pour améliorer l'utilisabilité de ces systèmes d'information en ligne et pour offrir de nouveaux services pertinents aux clients a été identifié par les entreprises. Les systèmes d'information doivent également pouvoir utiliser les ressources distribués et/ou des web services. Ainsi, la réutilisation et le partage des ressources nécessitent d'assurer l'interopérabilité et d'utiliser des standards (par exemple W3 C et IEEE), aussi bien que l'automatisation.

Actuellement, un système d'information est présenté aux utilisateurs comme un site web qui réutilise un ensemble de web services et est constitué de la composition des web services – comme par exemple dans les banques. Les web services sont fournis par différentes entreprises. Il est également nécessaire de réutiliser ces web services dans des contextes différents pour améliorer la réutilisabilité de ces web services et donc diminuer le coût de

maintenance. Demain, ces services de Web devront être adaptatifs, interopérables et flexibles pour assurer les différents besoins utilisateur et les différents contextes de réutilisation. Les nouveaux systèmes d'information sera adaptative, contextuelle, interopérable et composée de plusieurs web services adaptatifs. En d'autres termes, ces systèmes d'information feront l'utilisation simultanée de différents systèmes « adaptatifs ».

Objectifs

L'objectif principal de ce projet est d'étudier le problème de la réutilisation et de la composition des systèmes d'information adaptatifs et de créer un groupe européen capable de résoudre ce problème. Pour atteindre cet objectif, quelques sous objectifs ont été identifiés :

- Étudier les modèles et les méthodologies courantes pour développer les systèmes d'information adaptatifs et identifier les questions clés.
- Définir quel type de « standards » devrait être proposé sur ces différents modèles, comment réutiliser, partager ou échanger certaines parties de ces modèles pour assurer l'interopérabilité et la composition.
- Définir quel type d'architecture serait nécessaire pour composer ces systèmes adaptatifs.
- Établir un plan pour soutenir la création de nouvelles applications adaptatives en composant des services adaptatifs existants et en partageant les contenus, les modèles utilisateur et les fonctionnalités adaptatives.

Chapitre 7. Curriculum VITAE

Garlatti Serge
ENST Bretagne
Département Informatique
Technopôle de Brest Iroise – CS 83818
TAMCIC¹⁰, UMR CNRS 2872
29238 BREST Cedex
Laboratoire TAMCIC, "Traitement Algorithmique et Matériel de la Communication, de l'Information et de la Connaissance", CNRS, UMR 2872.

7.1 Formations et qualifications

- 1974 : Baccalauréat E, Lycée technique de Kérichen, Brest
- 1978 : Diplôme d'Ingénieur en fabrication électronique de l'Ecole Nationale d'Ingénieurs de Brest.
- 1980 : DEA d'informatique de l'université de Rennes 1.

- 1990 : Thèse en informatique de l'université de Rennes 1, sujet de la thèse : « Exception à l'héritage et logiques non monotones ». Mention très honorable.

7.2 Expérience professionnelle

- 1980 - 1983 : Thomson CSF, CEBM (Septembre 1980, Octobre 1983)
 - Ingénieur d'études, laboratoire Traitement et Transmission du Signal : conception d'un multiplexeur multiservices pour des liaisons troposphériques
- 1983 - 2005 : ENST Bretagne
 - 1983 - 1986 : Chargé d'enseignement recherche au département informatique et réseaux.
 - 1986 – 1992 : Maître de conférence au département informatique et réseaux
 - 1993 – 2000 : Maître de conférence au département Intelligence Artificielle et Sciences Cognitives.
 - 2001 – 2003 Directeur d'études au département Intelligence Artificielle et Sciences Cognitives.
 - 2003 – 2005 Directeur d'études au département Logiques des Usages, Sciences Sociales et de l'Information.
 - 2006 Directeur d'études au département Informatique.

7.3 Activités d'enseignement

Les activités d'enseignement ont principalement été orientées vers l'Intelligence artificielle, la représentation de connaissances et le raisonnement, les logiques standards et non standards, les systèmes d'informations adaptatifs et sémantiques.

7.3.1 Enseignements dispensés (principalement à l'ENST Bretagne)

1^{ère} Année tronc commun

- Module ISI 102, Logique, calculabilité, complexité, travaux dirigés de 1998 à 2001.

2^{ème} Année tronc commun

- Module ISI 201, Les bases de données, travaux dirigés de 1998 à 1999.

3^{ème} année et DEA (MIASH et IHM)

- Cours Prolog, (responsabilité, évolution, coordination, cours et Travaux pratiques) de 1986 à 1991 (aussi en mastère Images et Intelligence artificielle).
- Langages et Ateliers logiciels (Responsabilité, évolution, coordination et cours, projets) de 1998 à 1999.
- Les Systèmes à base de connaissances (Responsabilité, évolution, coordination et cours, de 1997 à 2002 (aussi en mastère Images et Intelligence artificielle).
- Résolution de problèmes, représentation de connaissances et raisonnement (DEA, Responsabilité, évolution, coordination et cours, de 1997 à 2005 (aussi en mastère Images et Intelligence artificielle).
- Les systèmes de recherche d'information (Option Informatique, Cours, de 1999 à 2002.
- Logiques classiques et non standards (DEA, responsabilité, évolution, coordination et cours) de 2001 à 2004.
- Les documents électroniques (évolution et cours) de 2001 à 2003.
- Les systèmes d'information orientés utilisateur (DEA, Responsabilité, évolution, coordination et cours) de 1997 à 2005.

Création et rédaction de supports de cours multimédia pour tous les modules de 3^{ème} année et mise à disposition sur Internet (<http://iasc.enst-bretagne.fr/~garlatti>).

7.3.2 Encadrements de projets

1^{ère} Année tronc commun

- Projet d'entreprendre, tutorat, de 2000 à 2005.
- Module IGS, ingénierie des grands systèmes, projet en pédagogie active, tutorat, de 2003 à 2005.

2^{ème} Année tronc commun

- Projet S4, encadrement de projets, de 2000 à 2005.

3^{ème} année et DEA

- Encadrement de Projets et de stages de l'option de troisième année CHMEST et de DEA.

7.3.3 Ingénierie pédagogique

Réalisation de polycopiés sur :

- Les systèmes formels et la logique classique
- Les logiques non standards
- Les représentations Objets

Réalisation de supports de cours :

- Domaine de l'intelligence artificielle
 - Introduction à l'I.A, les représentations Objet, les règles, le calcul des prédicats, la démonstration automatique de théorèmes et Prolog, les contraintes, le raisonnement
- Systèmes d'information orientés utilisateurs et documents électroniques
 - Les documents virtuels adaptatifs, les hypermédias adaptatifs, les systèmes hypermédia, les métadonnées, le web sémantique, les standards du W3C

7.3.4 Activités d'organisation

- Responsable du laboratoire « Atelier Logiciel » pour l'option de 3ème année de 1985 à 1989.
- Responsable des échanges d'étudiants et de la coopération académique avec l'Université du Sussex de 1995 à 1999.
- Responsable du Mastère Images et Intelligence Artificielle de 1989 à 1998, (organisation, évolution, emploi du temps, coordination avec le département ITI de l'ENST Bretagne et l'Institut d'Informatique Industrielle de Brest, interface avec les élèves).
- Responsable pédagogique de la filière Communications Multimodales et Coopération Homme/Machine du DEA Mathématiques, Informatique et Applications aux Sciences Humaines (organisation, emploi du temps, interfaces avec les élèves) de 1998 à 2002 et organisation de séminaires pour ce DEA.
- Responsable de l'option de 3ème année Coopération Homme Machine et Ergonomie dans les Services de Télécoms (organisation, emploi du temps, interfaces avec les élèves) de 2000 à 2002.
- Comité de pilotage pour la définition du projet de développement de 2ème année ou projet S3.

7.3.5 Activités de recherche

7.3.6 Domaines et axes de recherche

Dans un premier temps, mon activité de recherche était principalement orientée vers la représentation de connaissances et le raisonnement (représentations objets, héritage multiple avec exceptions et raisonnement et logique non monotone). Dans une perspective pluridisciplinaire, les relations de cette thématique avec certaines théories cognitives de la catégorisation (théorie du prototype) ont été étudiées dans le cadre d'un GDR de sciences cognitives (GDR 957).

Cette première thématique a été étendue vers la conception de systèmes d'information sémantiques fondés sur des représentations de connaissances formelles et ceci toujours dans une perspective pluridisciplinaire.

Approche proposée

De nos jours, les systèmes d'information (SI) sont conçus pour être accessibles sur Internet (hypermédia), possèdent une architecture distribuée, accèdent à une masse de ressources hétérogènes qu'il s'agit de réutiliser, assurent une adaptation / personnalisation en fonction des besoins utilisateur et permettent la coopération entre homme/machine et/ou homme/homme médiatisée par des machines au travers du web. Un système d'information adaptatif peut proposer différents contenus, différentes organisations et différentes présentations en fonction des besoins utilisateur. La conception de ces systèmes pose les problèmes suivants : une augmentation de la complexité liée à la masse des ressources disponibles et leur hétérogénéité qui pose des problèmes de recherche d'information, de réutilisation, de partage et d'échange de ces ressources, d'adaptation du système aux besoins utilisateur, de méthodes facilitant la maintenance et la conception de ce type de système – par exemple, un SI par utilisateur n'est pas envisageable. En d'autres termes, il est nécessaire d'automatiser le plus possible la création de ces SI en les adaptant aux utilisateurs.

L'objet principal de cette recherche est de définir une classe de méthodologies adaptées à la conception de systèmes d'information sémantiques et adaptatifs assurant une distribution du rôle de l'interprétation et de l'action entre l'homme et la machine. Ces systèmes d'information sont appelés sémantiques parce que leurs contenu et organisation sont fondés sur une modélisation du sens par des représentations de connaissances formelles adaptées aux besoins utilisateur. L'approche proposée est fondée sur :

- Les hypermédias adaptatifs : cadre conceptuel pour l'adaptation.
- Les documents virtuels : méthodologie permettant l'automatisation de la création de SI et facilitant la maintenance de systèmes d'information réutilisant des ressources distribuées.

- Le web sémantique et ses standards et la représentation des connaissances : association du sens aux ressources pour améliorer la recherche d'information et permettre la réutilisation de ressources et une automatisation de la création des SI.
- Des usages qui déterminent un contexte d'interprétation et d'action dans une communauté de pratiques particulière.

Les hypermédias adaptatifs fournissent un cadre conceptuel intéressant pour l'adaptation des systèmes d'information. Un hypermédia adaptatif est un hypertexte ou hypermédia qui adapte le contenu, la navigation et la présentation en fonction d'un modèle utilisateur. Les modèles utilisateur, issus de l'intelligence artificielle, prennent notamment en compte les connaissances, les tâches, les préférences et les compétences d'un utilisateur ou d'un groupe d'utilisateurs.

Un système d'information sémantique et adaptatif peut être vu comme un document virtuel adaptatif. Les documents virtuels sont des documents qui n'ont pas d'état persistant et pour lesquels chaque instance – document réel - est générée dynamiquement en fonction de la demande utilisateur. En d'autres termes, un document virtuel adaptatif possède un moteur de composition adaptatif qui sélectionne, organise et assemble dynamiquement des ressources multimédia pertinentes. Le document réel est généré dynamiquement à partir de modèles sémantiques – ontologies – et d'une spécification. Une spécification est un paramètre du moteur de composition qui peut être capable de définir au plus les critères de sélection, d'organisation, d'assemblage – présentation - et d'adaptation. Les documents virtuels apportent des méthodes et des techniques qui facilitent la création et la maintenance de systèmes d'information.

La recherche d'information par mots clés sur Internet possède de fortes limitations en termes de précision et de rappel. Par contre, les moteurs de recherche intelligents, fondés sur l'initiative du web sémantique - des métadonnées sémantiques, les ontologies correspondantes et des moteurs d'inférence - surmontent ces limitations en associant du sens aux ressources. Les métadonnées se substituent aux ressources et définissent un contexte de réutilisation de celles-ci. Le web sémantique améliore la qualité de la recherche d'information et autorise donc le partage, la réutilisation des ressources. Il permet ainsi de rendre le contenu du web interprétable par les machines. Il devient donc possible d'automatiser et/ou de semi automatiser la création des SI.

L'approche proposée s'appuie sur des pratiques : il s'agit de composer un document réel pour un utilisateur ayant des objectifs précis dans un contexte particulier : une communauté de pratiques. Les connaissances, compétences et savoir-faire explicites issus de ces communautés de pratiques permettent d'associer du sens.

Un Document Virtuel Adaptatif est donc composé d'un ensemble de ressources, d'ontologies, d'un schéma de métadonnées ainsi que d'un moteur de composition sémantique qui a pour but de sélectionner les ressources pertinentes, de les organiser et assembler en fonction d'une spécification de l'auteur ou du concepteur, et finalement d'adapter certains aspects visibles du document réel fourni à l'utilisateur.

Cette recherche s'articule avec des projets concrets, dans une approche pluridisciplinaire qui permet de définir un contexte d'interprétation et d'action dans une communauté de pratiques particulière.

7.3.6.1 Les enjeux

Dans le contexte de ma recherche qui s'inscrit dans celles du web sémantique – W3C, des documents virtuels et des hypermédias adaptatifs, parmi les enjeux les plus importants, on peut citer les suivants :

La description du sens, la distribution de l'interprétation et de l'action entre l'homme et la machine et l'interopérabilité des ressources au niveau sémantique.

- Associer du sens aux ressources peut ne pas avoir de limite : on peut imaginer que plus on décrit le sens à des niveaux sémantiques à la fois profonds et étendus plus on pourra améliorer la qualité de la recherche d'information, et donc plus on pourra automatiser certaines tâches. Il est bien évident qu'il y a une sorte d'optimum et/ou de ratio coût/gain acceptable dans chaque situation qui fixe les limites de la description du sens. Il y a malgré tout une limite due aux types d'activités mis en œuvre : certaines activités nécessitant une réelle expertise humaine ne sont pas transposables en machine. Il n'est pas possible d'acquérir les connaissances correspondantes.
- La répartition de l'interprétation et de l'action entre l'homme et la machine est directement liée à celle de l'attribution du sens. Dans une certaine mesure, fixer les limites de l'attribution du sens, c'est aussi fixer la répartition de l'interprétation et de l'action puisque les descriptions sémantiques ont pour objet de permettre à la machine d'automatiser et/ou de semi-automatiser certaines activités. Cette description du sens est réalisée à partir d'un schéma de métadonnées et des ontologies correspondantes qui assurent l'interopérabilité au niveau sémantique.
- Le moteur de composition sélectionne, organise et assemble des ressources hétérogènes et distribuées qui sont décrites par des métadonnées sémantiques et des ontologies. L'un des problèmes fondamentaux est de s'assurer que le document réel produit à partir de ces ressources soit cohérent et compréhensible par l'utilisateur. En d'autres termes, il faut que chaque ressource soit à la bonne place dans la bonne organisation et au bon moment. Il pose le problème de la qualité et de la précision de la recherche d'information – sélection - et de l'organisation pour un utilisateur donné. Ce problème est bien évidemment lié aux deux précédents la description du sens et la répartition de l'interprétation et de l'action entre l'homme et la machine puisqu'un moteur de composition a pour objet de prendre en charge les tâches machine de cette distribution.

Les propriétés de flexibilité et d'extensibilité d'un moteur de composition

- Flexible : plus les mécanismes du moteur de composition sont paramétrables, plus il est facile de maintenir et de concevoir de nouveaux documents virtuels. De nouveaux

documents réels peuvent être générés dès que de nouveaux paramètres de la sélection, de l'adaptation, de l'organisation et de l'assemblage sont définis et/ou modifiés.

- Extensible : à tout document virtuel sont associées des propriétés fondamentales qui caractérisent le type des documents réels produits. Un moteur de composition met en œuvre ces caractéristiques. Un moteur de composition n'est réutilisable que si ces propriétés fondamentales sont adéquates aux domaines d'application envisagés. En d'autres mots, si ces principes sont compatibles avec la conception d'un type de document virtuel particulier, on peut le réutiliser. Il est bien évident que ces principes fondamentaux sont aussi une limitation à la réutilisation. La conception du moteur de composition doit donc faciliter l'extension de ces propriétés.
- Organisation et acquisition des connaissances : comme le SI est fondé sur des connaissances, il est nécessaire d'avoir une méthodologie d'acquisition et de structuration des connaissances qui permettent une acquisition et une mise à jour aisée de celles-ci.

7.3.6.2 Quelques résultats

Ces recherches se sont concrétisées par la réalisation d'un moteur de composition sémantique appelé SCARCE – SemantiC & Adaptive Retrieval and Composition Engine) et une méthodologie de conception de systèmes d'information sémantiques et adaptatifs dont les propriétés principales sont les suivantes :

- Les connaissances explicites des communautés de pratiques sont représentées par des ontologies, des schémas de métadonnées sémantiques et de la logique. Elles servent de fondement aux modèles proposés. Ces connaissances permettent aussi d'assurer la cohérence des résultats et de fournir une méthodologie de conception de tels systèmes.
- Les mécanismes de sélection, d'organisation et d'adaptation peuvent être spécifiés au niveau sémantique par des paramètres déclaratifs issus des connaissances explicites de la communauté de pratiques pour assurer une flexibilité maximale du moteur de composition.
- Les ontologies sont organisées en quatre catégories l'une regroupant le modèle de document et le modèle d'adaptation, une ontologie du domaine, une ontologie des utilisateurs et une autre pour le schéma de métadonnées. Elles sont faiblement couplées et donc aisément modifiables. Une grande partie de celles-ci sont des paramètres du moteur de composition et peuvent donc changer et/ou mise à jour pour la maintenance et/ou la création de nouvelles applications.

Ce moteur de composition et la méthodologie correspondante ont été appliqués dans différents projets : ICCARS, CANDLE, KMP et Modales. Cette approche a aussi été reprise et étendue à la composition de services dans le projet CARISM et notamment dans le cadre de la thèse de O. Kassem Zein intitulée : « indexation, recherche et composition de services distribués ». Tous ces projets concrets ont permis d'avoir un retour d'expérience et ainsi d'affiner notre approche, mais aussi de l'étendre.

La société Atlantide qui a été impliquée dans les Projets SWAN et ICCARS, a repris cette approche dans les projets KNOSSOS et IMPERTINENTES. Ce transfert technologique leur permet maintenant d'avoir des offres dédiées à des SI fondés sur le sémantique Web.

7.3.6.3 Quelques perspectives

Parmi les perspectives actuelles, les trois axes principaux sont les suivants : i) la description du sens, la distribution de l'interprétation et de l'action entre l'homme et la machine et l'interopérabilité des ressources au niveau sémantique, ii) les propriétés de flexibilité et d'extensibilité du moteur de composition, iii) l'approche actuelle ne permet que la réutilisation de ressources statiques. Il serait intéressant d'étendre cette approche pour la réutilisation de services web et/ou de « composants actifs ».

Ces recherches ont apportés des réponses à ces questions, mais sans les épuiser. Il reste encore de nombreux aspects à améliorer ou à investir. Parmi ceux-ci on peut citer les suivants :

- L'adaptation est l'une des composantes de l'attribution du sens de la distribution de l'interprétation et de l'action qui est traité à un niveau sémantique. Il faudrait gérer les différentes stratégies d'adaptation à un niveau d'abstraction suffisant pour qu'elles soient directement interprétables par un utilisateur/auteur/concepteur. En effet, traiter uniquement de la présence/ absence d'hyperliens pour la navigation adaptative est un mécanisme de bas niveau. Il serait nécessaire de passer d'un modèle d'adaptation « simple » à un modèle d'adaptation et d'interaction déclaratif : l'adaptation peut être vue comme un quadruplet : ressources, modèle utilisateur, observateur et règles d'adaptation. Les règles d'adaptation comparent une description des ressources – métadonnées – au modèle utilisateur pour décider de leurs pertinences - un raffinement de la sélection. Avant d'appliquer les règles d'adaptation, l'observateur a enregistré certaines propriétés du comportement de l'utilisateur et s'en sert pour modifier son modèle. Un modèle d'interaction déclaratif est donc un modèle qui permet l'explicitation formelle de la boucle : Interaction (de l'utilisateur), Observation (par le système et mise à jour du modèle utilisateur) et Adaptation (application des règles) en relation avec le modèle utilisateur et le schéma de métadonnées. Cette explicitation doit faire directement références aux connaissances explicites de la communauté de pratiques et s'exprimer à niveau d'abstraction suffisant pour lui donner du sens.
- Etendre SCARCE avec un modèle de tâches hiérarchiques permettrait de rendre ce moteur de composition utilisable dans de plus nombreuses applications. Ceci est l'un des résultats des retours d'expérience des différents projets.
- Pour la réutilisation des services web, il est nécessaire d'intégrer la composition de services et de ressources – statiques – dans un unique environnement afin de pouvoir réutiliser d'autres web services et/ou des « composants actifs » - ressources actives. Pour des services web qui sont adaptatifs, le problème est plus complexe. Passer d'un moteur SCARCE centralisé à une architecture distribuée permettant la réutilisation et le partage de SI adaptatifs. Mais la réutilisation des services web adaptatifs reste encore un verrou technologique. De nombreux débats ont eu lieu dans la communauté

Hypermédia Adaptatifs sur ce sujet, mais sans solution probante pour l'instant. On peut ré-exprimer le problème de la manière suivante : comment doit-on décomposer un système d'information en services web – problème de granularité - pour assurer la réutilisation de services adaptatifs ? En effet, si l'on désire réutiliser un service web adaptatif, il est nécessaire de « connaître » et/ou « disposer » de l'information et de connaissances sur les différents éléments du quadruplet « ressources, modèle utilisateur, observateur, règles d'adaptation » pour concevoir un environnement d'apprentissage distribué, fondé sur des services web. Il s'agira donc de trouver le bon niveau de granularité des web services pour assurer cette réutilisation.

7.3.7 Thèses encadrées ou dirigées

Les principaux chercheurs encadrés sont les suivants :

- M. Victor Hugo Zaldivar Carillo (Taux encadrement : 40%) (1992-1995)
Sujet : Contributions à la formalisation de la notion de contexte ; le concept de "théorie" en représentation des connaissances.
- Melle Elisabeth Montabord (1993-1996, Taux encadrement : 100%)
Sujet : Base de connaissances, hypermédia et contexte,
- M. Nicolas Géhier (1995 -1998, Taux encadrement 100%)
Sujet : Classification conceptuelle incrémentale pour le raisonnement à base de cas,
- M. Sébastien Iksal (2000 – 2002, Taux encadrement : 100%)
Sujet : Spécification déclarative et composition sémantique pour des documents virtuels personnalisables,
- M. Jean-Louis Tetchueng, (2005, Taux encadrement : 100%)
Sujet : Modélisation du savoir-faire des enseignants pour la conception d'un EIAH adaptatif.

7.3.8 Stages de DEA encadrés ou dirigés

D'autre part, des stages de DEA ont été effectués sur les différents projets, notamment :

- M. Franck Magron : le raisonnement spatial pour la labellisation de structures anatomiques du cerveau ; (1997, Taux encadrement 100%)
- Melle Maggy Jobbard : Approche de la coopération homme/machine dans les systèmes hypermédia ; (1998, Taux encadrement 100%)
- M. Ugo Lobies : Hypermédia adaptatif d'aide à la navigation maritime ; (1998, Taux encadrement 100%)

- Sébastien Iksal : Coopération Homme Machine pour des Hypermédias Adaptatifs à l'aide d'une Modélisation Utilisateur ; (1999, Taux encadrement 100%)
- Benjamin Buffereau : Etude et réalisation de dossiers thématiques : un modèle de document virtuel personnalisable basé sur les standards du Web (2001, Taux encadrement 100%)
- Mme PHAM Thi Ngoc Diem : E-learning et résolution de problèmes (2003, Taux encadrement 100%)
- M. Jean-Louis Tetchueng : Intégration des DVP et de la pédagogie par problème pour la construction d'EIAH actifs et adaptatifs (2004, Taux encadrement 100%)

7.3.9 Administration et responsabilités collectives

- Co-responsable du groupe "Catégories, concepts et systèmes symboliques" dans le cadre du GDR "Sciences Cognitives de Paris" (C.N.R.S.) avec S. Desprès de l'université Paris VI de 1991 à 1994 et organisation d'un séminaire dans le cadre du GDR 957 "Sciences Cognitives de Paris" (Mars 94).
- Membres du comité de pilotage du réseau thématique pluridisciplinaire RTP 39 : « Apprentissage éducation et formation » depuis 2003, (<http://www-rtp39.imag.fr/>)
- Projet "Catégories, concepts et systèmes symboliques", GDR 957 "Sciences Cognitives de Paris" de 1991 à 1994.
- Projet " Analogie et Similitude", GDR 957 "Sciences Cognitives de Paris" à partir de Septembre 94.
- Projet PRC IA groupe "Classification et objets" depuis Septembre 1993.
- Participation à des Actions Spécifiques du CNRS, AS « Sémantique Web » du réseau thématique pluridisciplinaire RTP 33 "Documents et contenu : création, indexation, navigation" et de l'AS « WebLearn » du RTP 39.

7.3.10 Organisation de congrès

- Congrès international « Langage et Modèle à Objet », LMO'97, président du comité d'organisation, recherche de financements, publication d'un livre chez Hermès, (<http://www-lmo97.enst-bretagne.fr/>), Budget : 38 K€
- Congrès DVP 2002, Président du comité scientifique, organisation, recherche de financements, publication des actes, (<http://iasc.enst-bretagne.fr/DVP2002/>), Budget : 28 K€

7.3.11 Participation à des comités scientifiques

Membre de comités scientifiques de congrès nationaux :

- Représentations Objets, RPO de 1992 à 1993,
- Second séminaire francophone du Web Sémantique Médical, WSM 04 à Rouen
- RFIA 2004 et 2006, membre du comité éditorial, domaine couvert : documents virtuels et documents électroniques
- EIAH 2005,

Membre de comités scientifiques de workshops internationaux

- AH2003, the Adaptive Hypertext and Hypermedia Workshop at WWW2003, UM2003 and HT'03
- Adaptive Hypermedia and Collaborative Web-based Systems (AHCW'04) workshop held in conjunction with the International Conference on Web Engineering.
- Workshop "Authoring Learner and Pedagogical Models in Adaptive Educational Hypermedia" in conjunction with AIED'05
- SW-WL'05 in conjunction with CAISE'05, Semantic Web for Web-based Learning

Membre de comités scientifiques de congrès internationaux :

- Systemics, Cybernetics and Informatics, SCI, de 1997 à 1999,
- Langage Modèles et Objets, LMO de 1994 à 1999,
- IEEE/WIC/ACM International Conference Web Intelligence 2004, 2005
- IADIS International Conference, WWW / Internet 2004-2005
- Adaptive Hypermedia International Conference, 2004, 2006, Steering Committee.

Relecture dans des revues

- Revue « Hommes, Inter'actions et communication numériques », numéro spécial « problème de conception des revues adaptatives en ligne », comité de sélection et relecture. 2004
- Web Intelligence and Agent Systems, An International Journal, relecteur, 2004.

- “Distance et Savoir”, 2005.
- New Review of Hypermedia and Multimedia on Adaptive Hypermedia and the Adaptive Web, relecteur, numéro spécial 2004.
- Revue STICEF, Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation, 2004-2005.
- Revue STICEF, Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation, numéro spécial « Que signifient les ontologies pour les EIAH ? », 2004-2005.

7.3.12 Conférence invités, séminaires et tutoriaux

- Invited speaker at the Stefan Banach international mathematical center of Warsaw, Semester on algebraic logic and its applications, Novembre 91, Varsovie
- Garlatti, S., Concepts et représentations Objets, 1994, Séminaire du réseau COGNICENTRE.
- Garlatti, S., Objets et classes polythétiques, 1994, séminaire PRC GDR IA : Objets et Classifications.
- Garlatti, S., Concept representation and object centered paradigm, 1995, COGS Seminar, University of Sussex.
- Garlatti, S., The ATLAS Project, 1995, Intelligent Tutoring System Seminar, University of Sussex.
- 40èmes Journées de l'A.P.H.O. 1996. Perros-Guirrec, conférence invitée.
- 43th Meeting of the european working group "Multicriteria Aid for Decisions". 1996. Brest, tutorial.
- Workshop on Medical Image Tutoring, AI-ED'99, Le Mans, conférence invitée.
- 5th European Summer School, EUNICE'99, Barcelone, conférence invitée.
- Séminaire interdisciplinaire “LA EDUCACIÓN POLÍTICA : TEORÍA Y ADAPTACIÓN A LAS NUEVAS TECNOLOGÍAS, Madrid, 5, 6 Juin 2003, conférence invitée.
- Club E-learning de l’ANVIE, E-Learning et Technologies, conférence invitée, Octobre 2004.

- Ecole d'été régionale en informatique organisée par l'Institut de la Francophonie de l'Informatique (IFI) à l'Université des pêches de NhaTrang, 13-17 Septembre, cours : « Systèmes d'information orientés utilisateurs », conférence invitée.
- Journées « Intelligence collective, partage et redistribution des savoirs » organisées par le LGI2P (Ecole des Mines d'Alès) et le CERIC (Université Montpellier III), Septembre 2004, Nîmes, conférence invitée,

7.3.13 Mobilité, séjour Sabbatique

- Séjour sabbatique de 6 mois à l'université du Sussex au département « School of Cognitive and Computing Sciences ». Ce séjour a eu pour objet de développer une activité de recherche sur les systèmes d'aide à la décision dans le domaine médical.

7.3.14 Participation à des jurys de thèse

- Nacéra Madani, sujet : « Etude de l'héritage des propriétés dans les réseaux sémantiques. Notion de réseau d'héritage légal », 25 Mai 1994, Université Paris XIII, (Rapporteur)
- Victor Hugo Zaldivar-Carillo, sujet : « Contributions à la formalisation de la notion de contexte : le concept de « théorie » dans la représentation des connaissances », Université Montpellier II, 18 décembre 1995.
- Elisabeth Montabord, sujet : « Base de connaissances, hypermédia et contexte », Université de Rennes I, 15 Mars 1996.
- Sylvie Larvor, sujet : « Automatisation des messages bancaires issus de l'international », EHESS, 1998
- Christophe Vaudry, sujet : « composition dynamique d'informations dans la communication homme-machine », Université Montpellier II, 16 décembre 2002.
- Sébastien Iksal, sujet : « Spécification déclarative et composition sémantique pour des documents virtuels personnalisables », EHESS, 4 décembre 2002.
- Halima Habieb-Mammar, sujet : « EDPHA : un environnement de développement et de présentation d'hyperdocuments adaptatifs », Institut national des sciences appliquées de Lyon, 10 septembre 2004.
- John-Freddy Duitama Munõz, sujet : « un modèle de composants éducatifs pour des cours adaptatifs sur le web », Institut National des Télécommunications, avec l'université d'Evry-Val d'Essonne, 25 Mars 2005.

7.3.15 Expertise de projet

- Netherlands Organisation for Scientific Research, Physical Sciences Council, research proposal about Adaptive Hypermedia.
- Ministère à la recherche et aux nouvelles technologies, ACI Masses de données.

7.3.16 Activités d'organisation et responsabilités internes à l'ENST Bretagne

- Responsable du groupe de recherche en Intelligence Artificielle et systèmes cognitifs de 1989 à 1991 au département Informatique et réseaux de l'ENST Bretagne
- Responsable du projet SCRIpTureS (Semantic Composition and Retrieval of Information, Pictures and Services) depuis 2003 qui est devenu un projet structurant – projet de recherche - du Groupement des Ecoles des Télécommunications (ENST Paris, ENST Bretagne, INT, etc.) (http://www.get-telecom.fr/fr_accueil.html)

7.4 Recherche contractuelle

La première période de recherche s'est focalisée sur la représentation de connaissances et le raisonnement (représentations objets, héritage multiple avec exceptions et raisonnement et logique non monotone) et a été principalement marquée par le projet "Categories, Concepts and Symbolic Systems. Puis, l'activité a évolué en trois étapes principales retraçant l'évolution de l'activité de recherche et des technologies au travers de projets :

9. Les tâches/objectifs des utilisateurs sont pris en compte de manière implicites - il n'y a pas de modèles utilisateur.
 - Projet ATLAS : système d'aide au diagnostic pour le traitement des épilepsies focales, différents modèles de domaine du cerveau (neuroanatomie, neurophysiologie, etc.)
 - Projet ATLAS-TUTOR : système d'aide au diagnostic pour la détection de lésion du cerveau – tuteur intelligent, détection d'erreurs et remédiations.
10. Les besoins utilisateur sont représentés explicitement à l'aide d'un modèle utilisateur, SI adaptatif fondé sur les notions de document virtuel et d'hypermédia adaptatif / personnalisable
 - Projet SWAN : système d'aide à la navigation en ligne pour les marins : modèle utilisateur et modèle de tâches hiérarchiques
11. Documents virtuels adaptatifs fondés sur le web sémantique, la réutilisation et le partage de ressources distribuées, des catégories d'ontologies et un moteur de composition flexible et réutilisable.

- Projet ICCARS : Composition sémantique et adaptative de dossiers thématiques et projet ERGO-WEB pour l'évaluation de l'expertise des journalistes et du poste de travail proposé.
- Projet CANDLE : Composition sémantique et adaptative de cours en ligne
- Projet KMP : Portail sémantique pour du « Knowledge Management »
- Projet MODALES : Composition sémantique et adaptative de cours en ligne fondé sur des pratiques et théories didactiques

7.4.1 Projet "Categories, Concepts and Symbolic Systems"

Date : 1991-1994

Type : Programme Cogniscience, GDR 957

Objet : projet pluridisciplinaire sur les notions de catégories, de concepts et de systèmes symboliques dont l'objet était d'établir des correspondances entre les théories de la catégorisation et du concept en psychologie cognitive et leur représentation en mathématiques, en intelligence artificielle et plus particulièrement dans les représentations Objet.

Financements : 20 K€

7.4.2 Projets ATLAS

Dates : 1994-1997

Objet : Ce projet avait pour but la conception d'un environnement hypermédia d'aide au traitement des épilepsies focales pour les tâches suivantes : l'interprétation d'images, diagnostic et préparation d'actes chirurgicaux. Il s'agissait de fournir au médecin des informations adaptées à sa spécialité (neurochirurgie, neuroanatomie, neurophysiologique, etc.) pour les tâches citées précédemment. Cet environnement proposait un couplage entre des outils de traitement d'images et de gestion des connaissances.

Résultats : un prototype d'environnement hypermédia, une méthodologie de conception de systèmes à base de connaissances fondée sur la notion de perspectives et d'héritage multiple. Cette dernière permet d'offrir à l'utilisateur une recherche d'informations adaptatives (par requête ou navigation) en fonction de sa spécialité.

Rôles : Chef de projet, encadrement d'une équipe au département IASC, choix de l'architecture logicielle et des composants, organisation des réunions et suivi du projet, gestion des finances et investissements.

Environnement : Y3, environnement de développement de systèmes à base de connaissances développé par SEMA GROUP et l'INRIA pour les connaissances de domaines, le modèle de

tâche, l'hypermédia contextuel et donc l'interface homme/machine, base de données O2 pour les données patients, outils de traitement et d'analyse d'image utilisés par le laboratoire SIM, système d'exploitation Sun OS.

Partenaires : Internes : M. I. Kanellos ; enseignant/chercheur, M. Eric Faure, Ingénieur de recherche, M^{lle} Elisabeth Montabord - thésard -, stagiaires de DEA : M. Nicolas Gauthier, M. Franck Magron, stagiaires de fin d'études : M. Mickael Edstam, ainsi que d'autres stagiaires externes et étudiants en projet. M. A. Lasquelléc, département Informatique, M. J.H. Yapi, Thésard ; Externes : Laboratoire SIM, Hôpital Pontchaillou Rennes, M. Bernard Gibaud, M. Christian Barillot, Prof. J.M. Scarabin.

Finances : 100 K€

7.4.3 Projet ATLAS-TUTOR

Dates : 1996-1997

Type : projet Alliance entre la France et l'Angleterre, en collaboration avec l'université du Sussex (Brighton), l'université de Birmingham, l'université De Monfort (Leicester), Institut de neuroradiologie de Londres.

Objet : Ce projet avait pour objet d'étudier l'apport de connaissances anatomiques, ainsi que des connaissances spatiales qualitatives en neuroanatomie pour la conception d'un système d'enseignement assisté par ordinateur. Il s'agissait d'aider des radiologistes à décrire et interpréter des images de résonance magnétique du cerveau pour leur formation et ainsi offrir une assistance au diagnostic. Le système MR Tutor développé au Sussex avait une connaissance médicale limitée, mais fournissait une interface multimédia pour l'élicitation des descriptions d'images par un utilisateur. Le but est donc d'étudier une représentation de connaissances permettant au système de répondre à des descriptions incorrectes d'images faites par des "étudiants" à différents niveaux d'expertise. Les experts radiologistes acquièrent une bonne compétence en intégrant des connaissances médicales et pratiques. Cette intégration est un processus actif qui nécessite : a) d'articuler des connaissances médicales ; b) de les réfléchir sur des études de cas ; c) d'établir des connexions entre ces connaissances médicales et les expériences pratiques. Dans notre cas, il s'agissait d'étudier les connexions entre des connaissances anatomiques et spatiales sur les structures du cerveau et des lésions, et la description et l'interprétation des images.

Résultat : Une proposition d'environnement multimédia alliant les apports des deux projets MR Tutor et Atlas et permettant d'adapter l'aide aux apprenants en fonction de leurs différentes catégories d'erreurs ; une représentation de connaissances anatomiques utilisant une représentation qualitative des connaissances spatiales, une analyse des erreurs des apprenants et un atlas informatisé du cerveau.

Rôle : Gestion du projet, des échanges et des financements avec les partenaires

Partenaires : Internes : M. Eric Faure, Ingénieurs de recherche, Prof. J-B. Barthélemy ; Externes : De Montfort University, Leicester : Prof. B.A. Teather, Prof. D. Teather, Mr. A.I. Direne. ; Neurology Institute of London : Prof. G.H. du Boulay.; University of Sussex : Prof B. du Boulay; University of Birmingham : Prof. M. Sharples, Mr. N. Jeffery.

Financements : 3 K€, financement des déplacements France, Angleterre.

7.4.4 Projet SWAN

Dates : 1998, 1999

Type : Conseil Régional, CUB, programme ITR

Objet : Ce projet avait pour objet d'étudier et de mettre en œuvre une méthodologie de conception d'un serveur Web ou système d'information en ligne adaptatif - hypermédia adaptatif. Différents usagers sont intéressés par différentes catégories d'informations et différents moyens de navigation et/ou de recherche d'informations. Ce serveur proposait des informations nautiques pour la navigation ou la préparation de celle-ci qui étaient adaptées aux tâches des navigateurs, à ces préférences et à sa catégorie - plaisancier ou professionnel, aux caractéristiques du navire, de la météo - état de la mer, direction et force du vent - et de la marée.

Résultats : Une méthodologie de conception de système d'information adaptatif fondée sur des connaissances, un prototype de site Web pour des navigateurs et une organisation modulaire des connaissances munie d'un modèle de domaine, d'un modèle utilisateur et d'un modèle de tâches hiérarchiques. La méthodologie et l'organisation des connaissances associées furent le premier pas vers la notion de documents virtuels personnalisables/adaptatifs et donc de la composition. Le modèle de domaine jouait le rôle de l'indexation des informations incluses dans le SI.

Rôles : Chef de projet, encadrement d'une équipe au département IASC, des stagiaires, coordination des deux équipes de développement dont trois ingénieurs à Atlantide, organisation des réunions, suivi du projet, gestion des finances et investissements.

Environnement : WebObjects environnement de développement Objet de serveur Web dynamique interagissant avec un serveur HTTP, Loom et Lispworks environnement de développement du système à bases de connaissance pour l'implantation du modèle de domaine, du modèle utilisateur et du modèle de tâches, Java pour la réalisation d'applet. L'architecture est fondée sur un modèle client/serveur.

Partenaires : Internes : M. B. Gourvennec, Stagiaires de DEA : M^{elle} M. Jobard, M. Ugo Lobbies, Ingénieur de développement : M. S. Iksal, 3 étudiants de troisième année en projet ; Externes : Atlantide : M. P. Kervella, M. M. Touin, M. P. Bossard, M. F. Peault. EPSHOM : M. J. Fichant, M. C. Jégo.

Financements : 38 k€

7.4.5 Projet ICCARS

Dates : 1999, 2002

Type : Ce projet s'intégrait au réseau régional de la recherche en technologies de l'information et de la communication (R3TIC), Môle Armoricaïn de Recherche sur la Société de l'Information (Brest, Lannion, Rennes, Vannes) et au Programme de Recherche d'Intérêt Régional, presse régionale et technologie de l'information et de la communication.

Objet : Ce projet avait pour but d'étudier un système d'aide à la création de dossiers thématiques adaptatifs / personnalisables publiables sur internet. Un dossier thématique est une synthèse faite par des journalistes sur un sujet particulier.

Résultats : Un moteur de composition sémantique appelé SCARCE – SemantiC & Adaptive Retrieval and Composition Engine) dont le schéma de métadonnées et les ontologies correspondantes sont issus des connaissances communes et explicites d'une communauté de pratiques pour assurer la cohérence des résultats et fournir ainsi une méthodologie de conception de tels systèmes. Les mécanismes de sélection, d'organisation et d'adaptation peuvent être spécifiés au niveau sémantique par des paramètres déclaratifs issus des connaissances explicites de la communauté de pratiques pour assurer une flexibilité maximale du moteur de composition. Les ontologies sont organisées en quatre catégories l'une regroupant le modèle de document et le modèle d'adaptation, une ontologie du domaine, une ontologie des utilisateurs et une autre pour le schéma de métadonnées. Elles sont faiblement couplées et donc aisément modifiables. Une grande partie de celles-ci sont des paramètres du moteur de composition et peuvent donc changer et/ou mise à jour pour la maintenance et/ou la création de nouvelles applications. Un poste de journaliste assisté par ordinateur permettant aux journalistes de réaliser la spécification du dossier thématique.

Rôles : Chef de projet, encadrement d'une équipe au département IASC (ingénieurs,, stagiaires, thésard), coordination des deux équipes de développement dont deux ingénieurs à Atlantide, organisation des réunions, suivi du projet, gestion des finances et investissements.

Environnement : Environnement de gestion de connaissances Ontobroker, OntoEdit, servlets, serveur Apache, RMI, Java, DAML+ OIL, RDF/RDFS, XML, XSLT

Partenaires : Internes : thésard : M. S. Iksal, stagiaires de DEA : M. B Buffereau, Ingénieur de développement : M. P. Tanguy, étudiants de troisième année en projet ; Externes : Atlantide : M. P. Kervella, M. D. Riou ; Télégramme : M. O. Cleach, M. R Le Cleach, M^{me} A. Lessard.

Financements : 240 K€, Conseil Régional, programme ITR et PRIRE Presse.

7.4.6 Projet ERGOWEB

Dates : 2000, 2001

Type : Projet Cognitique 2000 sur deux ans

Objet : Le World Wide Web sur Internet devient un moyen extrêmement puissant de stocker et de diffuser de l'information. Pourtant, de nombreuses études attestent de la faible utilisabilité de ce média, voire de sa faible utilité : la majorité des utilisateurs n'y trouvent pas ce qu'ils y cherchent ; quand ils trouvent, l'information est souvent difficile à traiter. L'objectif du projet est d'étudier les processus cognitifs à l'œuvre dans le traitement de ce type de support. Nous voulons en particulier étudier : (a) les effets des nouveaux formats de présentation (intégration texte / image / vidéo / son) sur la compréhension ; (b) l'élaboration et la mise en œuvre de stratégies de recherche d'information, dans une perspective développementale (comment ces stratégies évoluent-elles avec l'expertise de l'utilisateur ?). Il est prévu (c) d'élaborer une méthode d'évaluation fondée sur l'interprétation des résultats de tests utilisateurs utilisant deux séries de variables (utilisabilité / utilité). Ces résultats pourront être exploités dans le contexte de la conception : fournir des méthodes, des recommandations ergonomiques pour la conception de sites Web. Dans le cadre de ce projet, des études liées à l'utilisabilité et l'utilité sont menées conjointement avec le projet Iccars.

Résultats : L'analyse des pratiques des journalistes a permis de confronter leur savoir faire à un modèle formel fondé sur la théorie des structures rhétoriques pour représenter les dossiers thématiques dans le système. Une analyse ergonomique du poste de journalisme assisté par ordinateur fondée sur ce modèle formel a été réalisée.

Environnement : Environnement de gestion de connaissances Ontobroker, OntoEdit, servlets, serveur Apache, RMI, Java, DAML+OIL, RDF/RDFS, XML, XSLT, Oracle.

Partenaires : Internes : thésard : M. S. Iksal, Ingénieur de développement : M. P. Tanguy, étudiants de troisième année en projet. ; Externes : Centre de Recherches en Psychologie, Cognition et Communication (EA 1285), Rennes, Université de Rennes 2 : M. J.E. Gombert, M. E. Jamet, F. Ganier, JC Coulet ; Pôle Création, Marketing et Acceptabilité de Services, France Telecom R&D – DIH, Lannion : M. G. Poulain ; Laboratoire Travail et Cognition, Unité Mixte de Recherche CNRS / Université (UMR 5551) Toulouse : M. A. Tricot

7.4.7 Projet CANDLE

Type : Projet européen IST sur trois ans (IST- 1999-11276)

Objet : Ce projet a pour objet l'utilisation d'Internet pour améliorer la qualité et réduire les coûts d'enseignement en Europe. Il s'agit d'utiliser le Web, la technologie multimédia et de permettre la coopération entre les universités et l'industrie dans la création, le partage et la réutilisation de matériaux pédagogiques. D'un point de vue recherche, il s'agit de mettre en œuvre des outils de création auteur et un système de production de cours adaptés aux apprenants et à différentes approches pédagogiques. Le savoir des auteurs est représenté à un niveau connaissance afin de le rendre partageable et réutilisable par les universités et les industries. Ce savoir faire est lié à l'approche pédagogique choisie.

Résultats : Les résultats sont en quelque sorte identiques à ceux de ICCARS puisque les deux projets ont été réalisés en même temps et ont été coordonnés tant du point de vue conceptuel que du point de vue développement.

Rôles : Responsable d'un « workpackage » (WP4 : information broker / applications), encadrement d'une équipe au département IASC (trois ingénieurs, des stagiaires et un thésard), coordination avec les différentes équipes du projet, organisation des réunions, suivi du projet, gestion des investissements.

Environnement : Environnement de gestion de connaissances Ontobroker, OntoEdit, servlets, serveur Apache, RMI, Java, DAML+OIL, RDF/RDFS, XML, XSLT, Oracle.

Partenaires : Internes : thésard : M. S. Iksal, Ingénieur de développement : M. E. Wiederhold, M. P. Wrona, P. Tanguy, étudiants de troisième année en projet ; Externes : British Telecom, Norwegian University of Science and Technology, Universität Karlsruhe, University of Twente, University College London, Politecnico Turino, Universität Stuttgart, Universitat Politècnica de Catalunya, Institute of Education, Institut National des Télécommunications, Suffolk College.

Financements : 252 K€ (2.3 M€ pour le projet dans sa globalité)

7.4.8 Projet KMP

Dates : 2004, 2005

Type : Projet RNRT

Objet : L'objectif du projet est de construire une solution innovante de "*Knowledge Management*" partagé entre différents acteurs, au sein d'une même communauté de pratique. Cette solution repose sur l'analyse, la conception, la réalisation et l'évaluation d'un prototype de service web de compétences, de type document virtuel personnalisable. Ce prototype sera une composante d'un portail web destiné à une communauté d'entreprises, d'institutionnels et d'organismes académiques intéressés par le domaine des télécommunications (communauté étudiée dans l'application : Telecom Valley). Ce projet est le fruit d'une réflexion menée dans le cadre du Laboratoire des Usages de Sophia Antipolis, dont la spécificité est de se focaliser sur l'observation des usages TIC dans le domaine de l'E-Business. Au regard des travaux existants, la contribution de ce projet est triple :

- Rendre compte des pratiques de Knowledge Management à l'échelle d'un réseau inter-firmes et inter-institutions ;
- Analyser les possibilités et opportunités de formalisation des pratiques d'échange et de création de connaissances inter-firmes, qui sont essentiellement informelles ;
- Concevoir, à partir de l'analyse et de l'anticipation des usages, le média le mieux approprié à ces pratiques.

Résultats : Un démonstrateur, KmP-Scarce fondé sur l'acquisition des pratiques des différentes entreprises, montrant l'intérêt des fonctionnalités de composition sémantique pour la gestion des compétences pour un portail de « Knowledge Management ». Une analyse fine des principes fondamentaux de SCARCE, suivie d'une étape de re-engineering de SCARCE

afin de s'affranchir le plus possible du moteur d'inférence et une proposition d'extension des principes fondamentaux de SCARCE pour pouvoir traiter ce type d'environnements. Il s'agit d'intégrer à SCARCE un modèle de tâches hiérarchiques. La tâche initialement prévue dans SCARCE, qui est la lecture d'un document dont la cohérence est le résultat du savoir-faire d'un auteur, sera traitée comme un tâche spécifique accessible dans le système.

Rôles : Sous-traitant d'un workpackage, encadrement d'une équipe au département IASC (1 ingénieur et des stagiaires), coordination avec les différentes équipes du projet, organisation des réunions, suivi du projet, gestion des investissements.

Environnement : SCARCE, Corèse

Partenaires : Université de Nice-Sophia Antipolis / CNRS (Latapses, UMR 6563), INRIA Sophia Antipolis, laboratoire Acacia, Association Telecom Valley, ENST / Département EGS

Financements : 100 K€

7.4.9 Projet MODALES

Dates : 2004-2008

Objet : Le projet de recherche MODALES (MODelling Didactic-based Active Learning Environment in Sciences) a pour objet l'analyse du savoir-faire des enseignants en didactique des sciences et en pédagogie ainsi que son application pour la mise en place de cours en ligne. Il s'appuie sur les compétences en didactique et pédagogie de l'IUFM, en psychologie cognitive du CRPCC et en conception d'EIAH « environnements informatiques pour l'apprentissage humain » vue comme des documents virtuels adaptatifs fondés sur l'architecture du « Semantic Web » par l'ENST Bretagne. Des situations concrètes d'apprentissage seront étudiées et mises en œuvre sur une plate-forme d'EIAH pour évaluer les résultats en terme de qualité d'apprentissage pour la population visée. La plate-forme développée dans le projet CANDLE servira de point de départ à ce projet (candle.enst-bretagne.fr) et sera étendue. Ce projet s'adresse à des élèves du primaire, du secondaire, à des étudiants ou à des professionnels de l'enseignement en situation de formation. Le thème choisi pour cette étude est l'atmosphère, (SVTE et Sciences Physiques) et peut-être abordé à différents niveaux (primaire, collège ou lycée, universitaire et formation des enseignants PE, PLC). L'objectif est de mettre l'apprenant dans un contexte d'apprentissage actif (pédagogie active ou apprentissage par problèmes ou problem-based learning) lors de la résolution de problèmes dans le cadre d'une formation en ligne. Il mettra l'apprenant en situation d'éprouver sa capacité à analyser un problème, à bâtir une stratégie de résolution, à opérer des choix par rapport au problème à résoudre : choix de la problématique, sélection des sources documentaires, modalité de sélection de ces sources. L'étude du savoir-faire des enseignants en termes de pédagogie active et de didactique des sciences servira de fondements pour la structuration des cours personnalisés. Ces principes de structuration et une sélection des ressources à disposition des enseignants permettront la génération dynamique des cours en utilisant les principes des documents virtuels adaptatifs. La structuration des cours et les contenus associés seront étudiés pour différentes catégories d'apprenants afin d'analyser les

variations interindividuelles et inter catégories en terme d'apprentissage. Il s'agira ici d'obtenir les paramètres pertinents qui permettront d'adapter les cours aux différentes catégories d'apprenants et dans une catégorie en fonction des individus. Il est bien évident que ces critères seront validés par une évaluation de la qualité d'apprentissage individuelle et par catégorie.

Rôles : Chef de projet, encadrement d'une équipe au département LUSI (stagiaires, thésard), coordination des trois équipes, organisation des réunions, suivi du projet, gestion des finances et investissements.

Environnement : SCARCE,

Partenaires : CREAD (Rennes2-IUFM de Bretagne), IUFM de Bretagne, centre de Brest et Rennes, Centre de Recherches en Psychologie, Cognition et Communication (EA 1285), Rennes, Université de Rennes 2

Financements : 300 K€

7.5 Publications

7.5.1 Revues

Garlatti, S. (1993). "Default logic models of inheritance systems with exceptions." *Fundamentae informatica*, special issue: algebraic logic and its applications 18(2, 3, 4): 129-150.

Gibaud, B., S. Garlatti, et al. (1998). "Computerized brain atlases as decision support systems: a methodological approach." *Artificial Intelligence in Medecine* 14: 83-100.

Garlatti, S. and M. Sharples (1998). "The Use of a Computerized Brain Atlas to Support Knowledge-based Training in Radiology." *Artificial intelligence in Medecine* 13: 181-205.

Garlatti, S. and Y. Prié (2005). "Adaptation et personnalisation dans le sémantique web." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004

Prié, Y. and S. Garlatti (2005). "Métadonnées et annotations dans le sémantique web." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.

7.5.2 Contribution à des ouvrages collectifs

Montabard, E., B. Gibaud, et al. (1993). *An hypermedia system to manage anatomical knowledge about brain. Computer Assisted Radiology*, Berlin, Springer-Verlag.

Garlatti, S., E. Montabard, et al., Eds. (1995). *A Methodological Approach for Object Knowledge Bases. Research and Development in Expert Systems XII*. Cambridge, SGES Publications.

Ducournau, R. et S. Garlatti, eds. *LMO'97 : Langages et modèles à objets*, 1997, Hermès.

- Gibaud, B., S. Garlatti, et al. (1997). Methodology for the design of digital brain atlases. Artificial Intelligence in Medicine (AIME 97). E. Keravnou, C. Garbay, R. Baud and J. Wyatt. Grenoble: 441-451.
- Garlatti, S. and S. Iksal (2000). Concept Filtering and Spatial Filtering in an Adaptive Information System. Adaptive Hypermedia and Adaptive Web-Based Systems. P. Brusilovsky, O. Stock and C. Strapparava. Trento, Springer Verlag: 315-318.
- Iksal, S., S. Garlatti, et al. (2001). Semantic composition of special reports on the Web: A cognitive approach. Hypertextes et Hypermédia H2PTM'01. P. Balpe, S. Lelu-Merviel, I. Saleh and J.-M. Laubin, Hermès: 363-378.
- Iksal, S. and S. Garlatti (2001). Revisiting and Versioning in Virtual Special Reports. Hypermedia: Openness, Structural Awareness and Adaptivity, Third Workshop on Adaptive Hypertext and Hypermedia. S. Reich, M. M. Tzagarakis and P. D. Bra. Århus, Denmark, Springer Verlag: 264-279.
- Garlatti, S. and M. Crampes, Eds. (2002). Documents Virtuels Personnalisables 2002. Brest, 10-11 Juillet, ENST Bretagne.
- Garlatti, S. and S. Iksal (2003). Declarative Specifications for Adaptive Hypermedia Based on a Semantic Web Approach. User Modeling. P. Brusilovsky, A. Corbett and F. d. Rosis. Pittsburg, Springer Verlag: 81-85.
- Garlatti, S. and S. Iksal (2004). A Flexible Composition Engine for Adaptive Web Sites. Adaptive Hypermedia and Adaptive Web-Based Systems, AH 2004. P. D. Bra and W. Nejdl, Springer Verlag. LNCS 3137: 115-125.
- Garlatti, S., S. Iksal, et al. (2004). SCARCE: an Adaptive Hypermedia Environment Based on Virtual Documents and Semantic Web. Adaptable and Adaptive Hypermedia Systems. S. Y. Chen and G. D. Magoulas., Idea Group Inc.: 206-224.

7.5.3 Actes de colloques internationaux avec comité de lecture

- Ducournau, R., S. Garlatti, et al. (1989). About exception handling for inheritance hierarchies. Workshop on inheritance hierarchies in knowledge representation and programming language, Viareggio.
- Ducournau, R., S. Garlatti, et al. (1990). Linear-time algorithms and non-monotonic reasoning in hierarchies. Cognitiva 90, Madrid, Afcet., p. 279-286.
- Garlatti, S. (1991). Inheritance of properties and default logic. 1st World Conference on the Fundamentals of Artificial Intelligence, Paris, CNRS, AFIA. P. 199-210.
- Garlatti, S. (1991). Inheritance of properties and default logic, Invitation Stefan Banach international mathematical center de Varsovie, Semester on algebraic logic and its applications. November.
- Montabord, E., S. Garlatti, et al. (1994). Interaction between a hypermedia and a knowledge base management system about the brain. neural networks and Expert Systems in Medecine and Healthcare, Plymouth, University of Plymouth: p. 136-145.
- Barillot, C., B. Gibaud, et al. (1994). An Information System to manage Anatomical Knowledge and Image Data about Brain. Visualisation in Biomedical Computing, Rochester, SPIE p. 424-434.

- Kanellos, I., S. Garlatti, et al. (1994). The ATLAS Project: A joint industrial and scientific development under the cognitive paradigm of information processing. First European Conference: Cognitive Science in Industry, Luxembourg. p. 451-466.
- Garlatti, S. (1996). Tutorial: Multimédia et systèmes d'aide à la décision en situation complexe. 43th meeting of the european working group "Multicriteria Aid for Decisions", Brest.
- Garlatti, S. (1997). Adaptive hypermedia for Decision Support Systems. 7th Euro Conference on Decision Support Systems and groupware multimedia electronic commerce, Brugges: p. 112.
- Gehier, N. and S. Garlatti (1998). Incremental Conceptual Clustering for Case-based Reasoning. IFCS 98, Rome.
- S. Garlatti, 1999, "Decision Support Systems as Digital Brain Atlases", Workshop on Medical Image Tutoring, AI-ED'99, Le Mans.: 10-22. (Invited paper).
- S. Garlatti, 1999 "Adaptive Web Server for On-line Information Systems", 5th European Summer School, EUNICE'99, Barcelone.: p. 197-204. (Invited paper).
- S. Garlatti, S. Iksal, et al., 1999, "Adaptive On-line Information System", Information System, Analysis and Synthesis, ISAS'99, Orlando, Floride: p. 1050-1057.
- S. Garlatti, S. Iksal, et al., 1999, "Adaptive On-line Information System by means of a Task Model and Spatial Views", Second Workshop on Adaptive Systems and user Modeling on the World Wide Web, Toronto (WWW8) and Banff (UM 99), Eindhoven University of Technology: p. 59-66.
- Iksal, S., S. Garlatti, et al. (1999). On-line multimedia Information System Adapted to Sailors, Human Centered Processes 10th Mini EURO Conference,, Brest, p. 431-436.
- Iksal, S. and S. Garlatti (2002). Adaptive Special Reports for On-line NewsPapers. Workshop Electronic Publishing, Adaptive Hypermedia (AH) 2002. S. Mizzaro and C. Tasso. Malaga, Espagne, Universidad de Malaga: p. 27-30.
- Garlatti, S. and S. Iksal, Eds. (2003). A Semantic Web Approach for Adaptive Hypermedia. Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems, Technische Universiteit Eindhoven, p. 5-14.
- Iksal, S. and S. garlatti (2004). Adaptive Web Information Systems: Architecture and Methodlogy for Reusing Content. AH 2004 workshops, EAW'04: Engineering the Adaptive Web, Eindhoven, Technische Universiteit Eindhoven: p. 36-45.
- Garlatti, S., J.-L. Tetchueng, et al. (2006). The Co-Design of Scenarios for a Didactic-based E-learning System viewed as an Adaptive Virtual Document. Satellite Workshop E-Learning, 2nd IEEE International Conference on Information & Communication Technologies : from Theory to Applications, Damascus, Syria
- Laubé, S., S. Garlatti, et al. (2006). Modelling Teachers Activities to Design Scenarios for a Didactic-based E-learning System viewed as an Adaptive Virtual Document. European. Conference on Education Research, Genova.

7.5.4 Colloques nationaux et colloques internationaux sans actes

- Garlatti, S. (1990). Exceptions à l'héritage et logiques non monotones, IFSIC, Université de Rennes I.

- Despres, S. and S. Garlatti (1991). Gestion du typique et de l'atypique dans les systèmes à héritage. Actes des Journées PRC IA, Plestin les Grèves.
- Garlatti, S. (1991). Atypicalité et héritage. Journées de Synthèse de l'atelier "Catégories, Concepts et Systèmes Symboliques, Brest.
- Garlatti, S. and I. Kanellos (1993). Les objets comme représentations des catégories naturelles ou artefactuelles. RePrésentation Objets, La Grande Motte.
- Gibaud, B., C. Barillot, et al. (1993). Brain atlases. Workshop AICOM 10, Bruxelles, AIM office.
- Garlatti, S., I. Kanellos, et al. (1994). YAKA : un système d'information intensionnel sur le cerveau humain. Informatique des Organisations et des systèmes d'information et de décision, Aix en provence, Inforsid. p. 323-338.
- Garlatti, S. (1994). Concepts et représentations Objets, Séminaire du réseau COGNICENTRE.
- Garlatti, S. (1994). Objets et classes polythétiques, PRC GDR IA: Objets et Classifications.
- Garlatti, S. (1995). The ATLAS Project, ITS Seminar, University of Sussex.
- Garlatti, S. (1995). Concept representation and object centered paradigm, COGS Seminar, University of Sussex.
- Garlatti, S. (1995). Systèmes à Base de Connaissances et Termes. Journées ERLAT-Glat, Brest.
- Garlatti, S. (1996). Les systèmes interactifs d'aide à la décision en situations complexes. 40e Journées de l'A.P.H.O., Perros-Guirrec.
- Garlatti, S., E. Montabond, et al. (1996). Base de connaissances intensionnelle avec perspectives et héritage multiple. Congrès de Reconnaissance des formes et d'Intelligence Artificielle, Rennes, Afia- Afcet. p. 495-504.
- Garlatti, S. and P. Kervella (1998). Les serveurs Web adaptatifs. Objet'98, Nantes.
- Garlatti, S. and S. Iksal (1999). Documents virtuels personnalisables pour des systèmes d'informations en ligne. Workshop sur les Documents Virtuels Personnalisables, IHM'99, Montpellier. p. 22-26.
- Garlatti, S. and I. Iksal (2000). Méthodologie de conception de documents électroniques adaptifs sur le Web. CIDE 2000, Lyon.
- Iksal, S. and S. Garlatti, Eds. (2001). Documents Virtuels et Composition Sémantique Une Architecture Fondée sur des Ontologies. NimeTic'01. Nimes, Ecole des Mines d'Alès. P. 91-96
- Iksal, S. and S. Garlatti (2002). Spécification déclarative pour des documents virtuels personnalisables. Documents Virtuels Personables 2002, Brest, 10-11 Juillet, ENST Bretagne. p. 127-140
- Garlatti, S. and Y. Prié (2003). Adaptation et personnalisation dans le sémantique web. Action spécifique 32 CNRS/STIC, Web Sémantique, rapport final. J. C. Charlet, C. Laublet and C. Reynaud: p. 79-89.
- Prié, Y. and S. Garlatti (2003). Métadonnées et annotations dans le sémantique web. Action spécifique 32 CNRS/STIC, Web Sémantique, rapport final. J. C. Charlet, C. Laublet and C. Reynaud: p. 25-32.

- Laubé, S., S. Garlatti, et al. (2005). Formations des maîtres en sciences : de la modélisation des situations didactiques comme fondement de la conception d'un EIAH adaptatif. EIAH 2005, Montpellier.
- Laubé, S., S. Garlatti, et al. (2005). Formations des maîtres en sciences: De la modélisation des situations didactiques comme fondement de la conception d'un Environnement Informatique pour l'Apprentissage Humain (EIAH) adaptatif. 4èmes rencontres de l'ARDIST, INRP.
- Laubé, S., Y. Kuster, et al. (2006). MODALES : Conception de scénario pour générer des EIAH adaptatifs pour la formation des maîtres en sciences expérimentales. Colloque " Scénarios, Formation d'enseignants : quels scénarios ? Quelles évaluations ?" IUFM de Versailles.
- Laubé, S., S. Garlatti, et al. (2006). Scénarios intégrant les TICE : les méthodologies et les cadres théoriques à l'œuvre dans la recherche MODALES. 8ème Biennale de l'Education, Colloque "Scénariser l'enseignement et l'apprentissage : une nouvelle compétence pour le praticien ?" INRP - ERTé E-Praxis, Paris.

Chapitre 8. Bibliographie Générale

- Akrich, M., M. Callob, et al. (1998). "A quoi tient le succès des innovations." Annales des Mines: 4-17.
- Aroyo, L., P. De Bra, et al. (2003). Embedding Information Retrieval in Adaptive Hypermedia: IR meets AHA! AH 2003: Workshop on Adaptive Hypermedia And Adaptive Web-Based Systems, Technische Universiteit Eindhoven.
- Attardi, G. and M. Simi (1985). Metalanguage and Reasoning Across Viewpoints. Advances in Artificial Intelligence. T. O'Shea, Elseviers Science Publishers B.V.: 413-422.
- Baget, J.-F., E. Canaud, et al. (2005). "Les langages du Web sémantique." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.
- Bajcsy, R. and S. Kovacic (1989). "Multiresolution Elastic Matching." CVGIP 46: 1-21.
- Balpe, J. (1997). Trois mythologies et un poète aveugle. Colloque "De la différence des arts", Paris, IRCAM.
- Barsalou, L. W. and D. L. Medin (1986). "Concepts: fixed definitions or dynamic context-dependent representations?" Cahiers de psychologie cognitive(6): 187-202.
- Berners-Lee, T. (1998). Why RDF model is different from the XML model. <http://www.w3.org/DesignIssues/RDF-XML.html>, W3C.
- Berners-Lee, T., J. Hendler, et al. (2001). "The semantic web." The Scientific American 5(284): 34-43.
- Bloom, F. E. (1990). Databases of brain information. Three-Dimensional Neuroimaging. A. W. Toga. New York, Raven Press: 273-295.

- Bobrow, D. G., R. Kaplan, et al. (1977). "GUS, a frame-driven dialog system." Artificial Intelligence 8: 155-173.
- Bobrow, D. G. and M. S. Stefik (1981). The LOOPS Manual, Xerox PARC.
- Bobrow, D. G. and T. Winograd (1977). "An overview of KRL, a Knowledge Representation Language." Cognitive Science 1: 3-46.
- Bohm, C., T. Greitz, et al. (1991). "Specification and selection of regions of interest (ROIs) in a computerized atlas." Journal of cerebral blood flow and metabolism(11): A64-A68.
- Bookstein, F. L. (1991). Thin-Plate Splines And The Atlas Problem For Biomedical Images. Information Processing in Medical Imaging. A. C. F. Colchester and D. J. Hawkes. Berlin, Springer-Verlag: 326-342.
- Borning, A. H. (1986). Classes versus Prototypes in Object-Oriented Languages. IEEE/ACM Fall Joint Conference.
- Boshuizen, H. P. A. and H. G. Schmidt (1993). "On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices." Cognitive science 16: 153-184.
- Boyle, C. and A. Encarnacion (1994). "Metadoc: an adaptive hypertext reading system." User Modeling and User-Adapted Interaction(4): 1-19.
- Bra, P. D. (1999). Design issues in adaptive hypermedia application development. Workshop on Adaptive Systems and User modeling on the World Wide Web, Toronto and Banff.
- Bra, P. D. and L. Calvi (1998). "Aha! an open adaptive hypermedia architecture." The new Review of Hypermedia and Multimedia 4: 115-139.
- Brachman, R. J. (1985). I lied about trees or defaults in knowledge representation. AI Magazine: 80-93.
- Brézillon, P. (2002). Hors du contexte, point de salut. Séminaire "Objets Communicants", Autrans, France.
- Brousseau, G. (1998). Théorie des situations didactiques. Grenoble.
- Brown, J. S., R. R. Burton, et al. (1982). Pedagogical, Natural Language, and Knowledge Engineering Techniques in SOPHIE I, II and III. Intelligent Tutoring Systems. D. H. Sleeman and J. S. Brown. London, Academic Press: 227-282.
- Brusilovsky, P. (1996). "Methods and techniques of adaptive hypermedia." User Modeling and User-Adapted Interaction 6(2-3): 87-129.
- Brusilovsky, P. (2001). "Adaptive Hypermedia." User Modeling and User-Adapted Interaction 11: 87-110.
- Brusilovsky, P., L. Eklund, et al. (1998). Web-based Education for all: A tool for development adaptive courseware. Seventh International World Wide Web Conference.
- Brusilovsky, P. and L. Pesin (1994). ISIS-Tutor: an adaptive hypertext learning environment. JCKBSE'94 Japanese-CIS Symposium on Knowledge-based software engineering, Pereslavl-Zalesski, Russia.
- Brusilovsky, P., L. Pesin, et al. (1993). Towards an adaptive hypermedia component for an intelligent learning environment. Human computer Interaction. L. J. Bass, J. Gornostaev and C. Unger. Berlin, Springer Verlag. 753: 348-358.
- Brusilovsky, P., E. Schwarz, et al. (1996). ELM-ART: An intelligent tutoring system on World Wide Web. Third International Conference on Intelligent Tutoring Systems, ITS-96, Montreal.

- Brusilovsky, P. and H. Su (2002). Adaptive Visualization Component of a distributed Web-base Adaptive Educational System. ITS 2002, Biarritz, Springer Verlag.
- Calfee, R. and M. J. Chambliss (1987). "The structural design features of large texts." Educational Psychologist 3-4(22): 357-378.
- Carré, B. and G. Comyn (1988). On Multiple Classification, Points of View and Object Evolution. Artificial Intelligence and Cognitive Sciences. J. Demongeot, T. Herve, V. Rialle and C. Roche, Manchester University Press: 49-62.
- Carré, B. and G. Comyn (1989). ROME, Multiple and evolutive object representation. Proceedings of the Fifth AIICSR conference, Czechoslovakia, North Holland.
- Carré, B., L. Dekker, et al. (1990). Multiple and Evolutive Representation in the ROME Language, Université de Lille.
- Carré, B. and J.-M. Geib (1990). The Point of View Notion for Multiple Inheritance. ECOOP/OOPSLA, ACM Sigplan Notices.
- Chaffin, R. (1981). "Context effects in the categorization task." The Journal of General Psychology(104): 293-302.
- Charlet, J. (2002). L'Ingénierie des connaissances : développements, résultats et perspectives pour la gestion des connaissances médicales. Informatique. Paris, Université Pierre et Marie Curie.
- Charlet, J., B. Bachimont, et al. (2005). "Ontologies pour le web sémantique." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.
- Charlet, J., P. Laublet, et al., Eds. (2003). Action spécifique 32 CNRS/STIC, Web Sémantique, rapport final.
- Charlet, J., P. Laublet, et al. (2005). "Le Web sémantique." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.
- Chevallard, Y. (1992). "Concepts fondamentaux de la didactique: perspectives apportées par une approche anthropologique." La Pensée sauvage, Grenoble, Recherches en didactique des mathématiques 12(1): 73-112.
- Chevallard, Y. (1999). "L'analyse des pratiques enseignantes en théorie anthropologique du didactique." La Pensée sauvage, Grenoble, Recherches en didactique des mathématiques 19(2): 221-226.
- Chevrin, V., J. Rouillard, et al. (2006). Multi-channel Adaptive contexte-sensitive intermediation : Reconciling Marketing goals and User Needs. Workshop MAC, Glasgow, Scotland, UK.
- Chouraqui, E. and P. Dugerdil (1988). Conflict solving in a frame-like multiple inheritance system. European conference on Artificial Intelligence.
- Christophides, V. (1998). Electronic Document Management Systems. <http://www.ics.forth.gr/~christop/>, UCH/FORTH.
- Clayton, B. (1985). ART, Automated Reasoning Tool, Inference Corporation, Los Angeles, California.
- Cohn, A. G. (1993). Modal and Non Modal Qualitative Spatial Logics. Workshop on Spatial and Temporal Reasoning at IJCAI93, Chambéry.
- Cohn, A. G. and N. M. Gotts (1994). A Theory of Spatial Regions with Undetermined Boundaries. Topological Foundations of Cognitive Science. C. Eschenbach, C. Habel and B. Smith.
- Cohn, A. G., D. A. Randell, et al. (1994). Taxonomies of Logically Defined Qualitative Spatial Relations. Formal Ontology in Conceptual Analysis and

- Knowledge Representation. N. Guarino and R. Poli, Kluwer.
- Cohn, A. G., D. A. Randell, et al. (1993). Qualitative Spatial Reasoning and Representation. QUARDET: Qualitative Reasoning and Decision Technologies, Barcelona.
- Coirier, P., D. Gaonac'h, et al. (1996). Psycholinguistique textuelle : Approche cognitive de la compréhension et de la production des textes., Puf.
- Cole, M. and Y. Engeström (1993). A cultural-historical approach to distributed cognition. Distributed cognition. Psychological and educational considerations. G. Salomon, Cambridge University Press: 1-46.
- Condamines, A. and P. Amsili (1993). Terminology between language and Knowledge: An Example of Terminological knowledge Base. Terminology and Knowledge Engineering, Cologne.
- Conklin, J. (1987). "Hypertext: An introduction and Survey." IEEE Computer.
- Conlan, O., D. Lewis, et al. (2003). Applying Adaptive Hypermedia Techniques to Semantic Web Service Composition. Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems in The Twelfth International World Wide Web Conference, Budapest, Hungary.
- Crampes, M., S. Ranwez, et al. (2000). Ontology-supported and ontology-driven conceptual navigation on the world wide web. Proceedings of Hypertext 2000, San Antonio, USA, ACM Press.
- Cui, Z., A. G. Cohn, et al. (1993). Qualitative and Topological Relationships in Spatial Databases. Lecture Notes in Computer Science No. 692. D. Abel and B. C. Ooi. Berlin, Springer Verlag: 296-315.
- Davis, H. E. (1987). VIEWS: multiple perspectives and structured objects in a knowledge representation language, MIT.
- Dekker, L. and B. Carré (1992). Multiple and Dynamic Representation of Frames with Points of View in FROME. RePrésentation Objets, La grande Motte.
- DeMichiel, L. G. and R. P. Gabriel (1987). The common Lisp Object System: an Overview. ECOOP.
- Denhière, G. and S. Baudet (1992). Lecture, compréhension de texte et science cognitive. Paris, PUF.
- Derycke, A., V. Chevrin, et al. (2005). Intermédiations Multicanales et Multimodales pour l'E-formation : l'architecture du projet Ubi-Learn. EIAH 2005, Montpellier.
- Derycke, A., F. Hoogstoel, et al. (2006). Infrastructure des environnements informatiques pour l'apprentissage distribué. à paraître. J. Labat and M. Grandbastien, Hermès Sciences Publications.
- Derycke, A., Y. Peter, et al. (2005). "Conception d'une Plate-forme pour la recherche en EIAH." STICEF 12.
- Dijk, T. A. V. and W. Kintsch (1983). Strategies of discourse comprehension. Orlando, Academic Press.
- Downs, R. M. and D. Stea, Eds. (1973). Image and Environnement: Cognitive mapping and Spatial behaviour. Chicago, Aldine.
- du Boulay, G. H., B. A. Teather, et al. (1994). Standard Terminology for MR Image Description. XV Symposium Neuroradiologicum, Kumamoto.
- Ducournau, R. and M. Habib (1989). "Un mécanisme d'héritage dans les langages à objets." Techniques et Sciences Informatiques 8(1): 42-62.

- Ducournau, R., M. Habib, et al. (1995). "L'héritage Multiple dans tous ses états." Technique Scientifique Informatique 14(3).
- Dugerdil, P. (1986). A propos des mécanismes d'héritage dans un langage orienté objet. CIIAM 86 Intelligence Artificielle, Marseille.
- Dugerdil, P. (1988). Contribution à l'étude de la représentation des connaissances fondée sur les objets, Université d'Aix Marseille II.
- Dugerdil, P. (1991). Inheritance mechanisms in the Objlog Langage: multiple selective and multiple vertical with points of view. Inheritance Hierarchies in Knowledge Representation and Programming Languages. M. Lenzerini, D. Nardi and M. Simi. Viarregio, John Wiley & Sons Ltd: 245-256.
- Duursma, C., O. Olsson, et al. (1993). Task model definition and task analysis process. Deliverable ESPRIT Project P5248 KADS-II/M5/VUB/RR/004/2.0. T. R. CK-VUB-04, Free University Brussels.
- Ehrlich, M.-F., A. Charles, et al. (1992). La superstructure dans les textes expositifs est-elle prise en charge lors de la sélection des informations importantes ? Le résumé de texte. M. Charolles and A. Petitjean: 183-202.
- Encarnação, L. M. (1995). "Adaptivity in graphical user interface: an experimental framework." Computers & graphics 19(6): 873-884.
- Englert, C. S., S. R. Stewart, et al. (1988). "Young writers' use of text structure in expository text generation." Journal of Educational Psychology(80): 143-151.
- Evans, A. C., D. L. Collins, et al. (1992). "An MRI-based stereotactic atlas from 250 young normal subjects." Soc. Neurosci. Abstr(18): 408.
- Evans, A. C., T. S. Marrett, et al. (1991). "Mri-Pet Correlative Analysis Using A Volume Of Interest (Voi) Atlas." J. of Cerebral Blood Flow and Metabolism 11(2): A69-A78.
- Falquet, G., L. Nerima, et al. (1999). Des documents virtuels pour lire les bases de données. Workshop sur les Documents Virtuels Personnalises, IHM'99,, Montpellier.
- Fayol, M. (1985). Le récit et sa construction. Neuchatel, Paris, Delachaux & Niestlé.
- Fayol, M. and J.-M. Monteil (1988). "The notion of "script" : from general to developmental and social psychology." Cahiers de Psychologie Cognitive(8): 335-361.
- Fensel, D., J. Angele, et al. (1999). On2broker: Semantic-Based Access to Information Sources at the WWW. World Conference on the WWW and Internet, WebNet 99, Honolulu, Hawaii, USA.
- Ferber, J. (1988). Coreferentiality: The key to an Intentional Theory of Object-Oriented Knowledge Representation. Artificial Intelligence and Cognitive Science. J. Demongeot, T. Herve, V. Rialle and C. Roche, Manchester University Press. Chap 6.
- Ferber, J. and P. Volle (1988). Using coreference in object oriented representations. European Conference of Artificial Intelligence.
- Fischer, G. (1990). "Communication requirements for cooperative problem solving." Informations systems 15(1): 21-36.
- Folman, S. and S. Gissi (1990). "Intercultural rhetorical differences in meaning construction." Communication & Cognition 23(1): 45-92.
- Fox, P. T., J. S. Perlmuter, et al. (1985). "A Stereotactic Method of Anatomical Localization for Positron Emission Tomography." J. of Computer Assisted Tomography 9(1): 141-153.

- Fraissé, S. (1997). Une approche systématique pour la conception de documents hypermédias. Hypertextes et Hypermédia. J. P. Balpe, A. Lelu, M. Nanrd and I. Saleh, Hermes: 249-260.
- Freksa, C. (1991). Qualitative spatial reasoning. Cognitive and Linguistic Aspects of Geographic Space. D. M. Mark and A. U. Frank. Dordrecht, Kluwer: 361-372.
- Friston, K. J., R. E. Passingham, et al. (1989). "Localization in PET images: Direct Fitting of the Intercommissural (AC-PC) Line." J. Cereb. Blood Flow Metab 9(1): 141-153.
- Garlatti, S. (1997). Adaptive hypermedia for Decision Support Systems. 7th Euro Conference on Decision Support Systems and groupware multimedia electronic commerce, Brugges.
- Garlatti, S. and S. Iksal (1999). Documents virtuels personnalisables pour des systèmes d'informations en ligne. Workshop sur les Documents Virtuels Personnalises, IHM'99,, Montpellier.
- Garlatti, S. and S. Iksal (2000). Concept Filtering and Spatial Filtering in an Adaptive Information System. Adaptive Hypermedia and Adaptive Web-Based Systems. P. Brusilovsky, O. Stock and C. Strapparava. Trento, Springer Verlag: 315-318.
- Garlatti, S. and S. Iksal (2004). A Flexible Composition Engine for Adaptive Web Sites. Adaptive Hypermedia and Adaptive Web-Based Systems, AH 2004. P. D. Bra and W. Nejdl, Springer Verlag. LNCS 3137: 115-125.
- Garlatti, S., S. Iksal, et al. (1999). Adaptive On-line Information System by means of a Task Model and Spatial Views. Second Workshop on Adaptive Systems and user Modeling on the World Wide Web, Toronto and Banff, Eindhoven University of Technology.
- Garlatti, S., S. Iksal, et al. (2004). SCARCE: an Adaptive Hypermedia Environment Based on Virtual Documents and Semantic Web. Adaptable and Adaptive Hypermedia Systems. S. Y. Chen and G. D. Magoulas., Idea Group Inc.: 206-224.
- Garlatti, S., I. Kanellos, et al. (1994). YAKA : un système d'information intensionnel sur le cerveau humain. Informatique des Organisations et des systèmes d'information et de décision, Aix en provence, Inforsid.
- Garlatti, S., E. Montabord, et al. (1995). Knowledge Base Design in Solving Complex Problems. Brest, Télécom Bretagne.
- Garlatti, S. and Y. Prié (2005). "Adaptation et personnalisation dans le sémantique web." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.
- Garlatti, S. and M. Sharples (1998). "The Use of a Computerized Brain Atlas to Support Knowledge-based Training in Radiology." Artificial intelligence in Medecine 13: 181-205.
- Garlatti, S., J.-L. Tetchueng, et al. (2006). The Co-Design of Scenarios for a Didactic-based E-learning System viewed as an Adaptive Virtual Document. Satellite Workshop E-Learning, 2nd IEEE International Conference on Information & Communication Technologies : from Theory to Applications, Damascus, Syria.
- Gee, J. C., C. Barillot, et al. (1994). Matching Structural Images of the Human Brain using Statistical and Geometrical Image Features. SPIE, Visualization in Biomedical Computing, SPIE.
- Geurts, J. P., S. Bocconi, et al. (2003). Towards ontology-driven discourse: from semantic graphs to multimedia presentations, Centrum voor Wiskunde en Informatica.
- Goldstein, I. P. and D. G. Bobrow (1980). Extending Object Oriented Programming in Smalltalk. The lisp Conference, Stanford University.

- Greenacre, M. J. (1993). Correspondence Analysis in Practice, Academic Press.
- Greitz, T., C. Bohm, et al. (1991). "A computerized brain atlas: construction, anatomical content, and some applications." Journal of Computer Assisted Tomography 15(1): 26-38.
- Gruber, T. R. (1993). Toward principles for the design of ontologies used for knowledge sharing. International Workshop on Formal Ontology, Kluwer Academic.
- Gruber, T. R., S. Vemuri, et al. (1996). Model-based Virtual Document Generation, Knowledge Systems Laboratory.
- Grunst, G. (1993). Adaptive hypermedia for support systems. Adaptive user interfaces: Principles and Practice. M. Schneider-Hufschmidt, T. Küme and U. Malinowski. Amsterdam, North-Holland: 269-283.
- Guha, R. V. (1990). Contexts in Cyc, Stanford University, Austin Texas.
- Guha, R. V. (1991). Contexts: A Formalization and Some Applications, Standford.
- Guha, R. V. and D. B. Lenat (1990). Cyc: A Midterm Report. AI Magazine. 3: 33-59.
- Hahn, W. V. and W. Wahlster (1983). Ham-ans: a german language dialog system for knowledge-based access to relational data bases. SIGART Newsletter.
- Hart, R. A. and G. T. Moore (1973). The development of spatial cognition: a review. Image and Environnement: Cognitive Mapping and Spatial Behaviour. R. M. Downs and D. Stea. Chicago, Aldine: 111-143.
- Henze, N. and W. Nejdl (2004). "A Logical Characterization of Adaptive Educational Hypermedia." New Review on Hypertext and Hypermedia (NRHM) 10(1): 77-113.
- Hernandez, D. (1993). Reasoning with qualitative representations: Exploiting the Structure of Space. QUARDET, III IMACS international Workshop on Qualitative Reasoning and decision technologies, Barcelona.
- Hernandez, D. (1994). Qualitative Representation of spatial Knowledge. Berlin, Springer verlag.
- Hernandez, D. and K. Zimmermann (1993). Default reasoning and the qualitative representation of spatial knowledge. defaults and prototypes - Non-monotonic Reasoning for Language knowledge Processing. H. e. al, Springer verlag.
- Höhne, K. H., M. Bomans, et al. (1992). "A volume based anatomical atlas." IEEE Comp. Graphics & Appl. 12(9): 72-78.
- Höök, K., J. Karlgren, et al. (1996). "A glass box approach to adaptive hypermedia." User Models and User Adapted Interaction(6).
- Iksal, S. (2002). Spécification Déclarative pour des documents virtuels personnalisables. IASC. Brest, ENST Bretagne.
- Iksal, S. and S. Garlatti (2001). Revisiting and Versioning in Virtual Special Reports. Hypermedia: Openness, Structural Awareness and Adaptivity, Third Workshop on Adaptive Hypertext and Hypermedia. S. Reich, M. M. Tzagarakis and P. D. Bra. Århus, Denmark, Springer Verlag: 264-279.
- Iksal, S. and S. Garlatti (2004). Adaptive Web Information Systems: Architecture and Methodlogy for Reusing Content. AH 2004 workshops, EAW'04: Engineering the Adaptive Web, Eindhoven, Technische Universiteit Eindhoven.
- Isaac, A., P. Courounet, et al. (2004). Un système d'annotation multiforme et communautaire de documents audiovisuels : OPALES. Journée sur les Modèles Documentaires de l'Audiovisuel, Semaine du Document Numérique, La Rochelle.
- Kaheneman, D. (1973). Attention and Effort. Englewood Cliffs, Prentice Hall.

- Kaplan, C., J. Fenwick, et al. (1993). "Adaptive hypermedia Navigation Based on User Goals and Context." User Models and User-Adaptive Interaction 3(3): 193-220.
- Kay, J. (1994). Lies, Damned Lies and Stereotypes: Pragmatic approximations of users. Conference on user modeling, Hyannis, MA.
- Kay, J. and R. Kummerfeld (1994). An individualised Course for the C programming Language. Second international WWW conference "Mosaic and the Web", Chicago, IL.
- Keyserlingk, D. G. v., K. Niemann, et al. (1988). "A quantitative approach to spatial variation of human cerebral sulci." Acta Anatomica(131): 127-131.
- Keyserlingk, D. G. v., K. Niemann, et al. (1985). "A new method in computer assisted imaging in neuroanatomy." Acta Anatomica(123): 240-246.
- Knudsen, J. L. (1988). Name collision in multiple classification hierarchies. ECOOP.
- Kobsa, A. (1993). User modeling: Recent Work, Prospects and Hazards. Adaptive User Interfaces: Principles and Practice. M. Schneider-Hufschmidt, T. Kühme and U. Malinowski. Amsterdam, North-Holland.
- Kobsa, A., D. Müller, et al. (1994). KN-AHS: an adaptive hypertext client of the user modeling system BGP-MS. 4th International Conference on User Modeling, Hyannis, MA.
- Kobsa, A. and W. Pohl (1995). "The user modeling shell system bgp-ms." User Modeling and User-adapted Interaction(4): 59-106.
- Kobsa, A. and W. Wahlster (1989). User Models in Dialog Systems, Springer Verlag.
- Laubé, S., S. Garlatti, et al. (2006). Scénarios intégrant les TICE : les méthodologies et les cadres théoriques à l'œuvre dans la recherche MODALES. 8ème Biennale de l'Education, Colloque "Scénariser l'enseignement et l'apprentissage : une nouvelle compétence pour le praticien ?" INRP - ERTÉ E-Praxis, Paris.
- Laubé, S., S. Garlatti, et al. (2006). Modelling Teachers Activities to Design Scenarios for a Didactic-based E-learning System viewed as an Adaptive Virtual Document. European. Conference on Education Research, Genova.
- Laubé, S., S. Garlatti, et al. (2005). Formations des maîtres en sciences : de la modélisation des situations didactiques comme fondement de la conception d'un EIAH adaptatif. EIAH 2005, Montpellier.
- Laubé, S., S. Garlatti, et al. (2005). Formations des maîtres en sciences: De la modélisation des situations didactiques comme fondement de la conception d'un Environnement Informatique pour l'Apprentissage Humain (EIAH) adaptatif. 4èmes rencontres de l'ARDIST, INRP.
- Laubé, S., Y. Kuster, et al. (2006). MODALES : Conception de scénario pour générer des EIAH adaptatifs pour la formation des maîtres en sciences expérimentales. Colloque "Scénarios, Formation d'enseignants : quels scénarios ? quelles évaluations ?" IUFM de Versailles.
- Lehmann, E. D., D. J. Hawkes, et al. (1991). "Computer aided interpretation of SPECT images of the brain using an MRI derived 3D neuro-anatomical atlas." Medical Informatics 16(2): 151-166.
- LTSC (1999). Learning Object Metadata Working draft document 3. <http://ltsc.ieee.org/doc/wg12/LOM-WD3.htm>, Learning Technology Standard Committee.

- Maedche, A., S. Staab, et al. (2001). Seal - a framework for developping semantic portals. BNCOD, British National Conference on Databases, Oxford, UK, Springer Verlag.
- Mann, W. C. and S. A. Thompson (1988). "Rhetorical Structure Theory: Toward a functional theory of text organization." Text 8(3): 243-281.
- Marino, O. (1989). Classification d'objets dans un modèle multi-points de vue, Laboratoire ARTEMIS/IMAG, INPG.
- Marino, O. (1991). Classification d'objets composites dans un système de représentation des connaissances multi-point de vue. RFIA.
- Marino, O. (1993). Raisonnement classificatoire dans une repr sentation objets multi-points de vue, Universit Joseph Fourier, Grenoble.
- Marino, O., F. Rechenmann, et al. (1990). Multiple perspectives and classification mechanism in object-oriented representation. European Conference on Artificial Intelligence.
- Mark, D. M. and A. U. Frank, Eds. (1991). Cognitive and Linguistic Aspects of Geographic Space. NATO Advanced Studies Institute. Dordrecht, Kluwer.
- Markowitz, J. A., J. T. Nutter, et al. (1992). "Beyond is-a and Part-Whole : More Semantic Network Links." Computers Math. Applic 23(6-9): 377-390.
- Martin, W. A. (1979). Descriptions and specialization of concepts. Artificial Intelligence, an MIT perspective. P. Winston and R. H. Brown, The MIT Press. 1: 377-419.
- Mathews, S. R. (1982). The impact of prior knowledge on accessibility and availability of information from prose. Discourse Processing. A. Flammer and W. Kintsch. Amsterdam, North-Holland.
- Mazziotta, J. C., A. W. Toga, et al. (1993). A Probabilistic Reference System for the Human Brain, Application to the Human Brain Project: Phase I.
- McNamara, T. P. (1986). "Mental Representation of Spatial Relations." Cognitive Psychology(18): 87-121.
- McTear, M. F. (1993). "User modelling for adaptive computer systems: a survey of recent developments." Artificial Intelligence Review 7: 157-184.
- Medin, D. L. and E. J. Shoben (1988). "Context and structure in conceptual combination." Cognitive psychology(20): 158-190.
- Meyer, B. (1988). Object-oriented software construction, Prentice Hall.
- Milosavljevic, M., F. Vitali, et al. (1999). Introduction of the virtual document workshop. Workshop on Virtual Document, Hypertext Functionality and the Web, Toronto.
- Minski, M. (1975). A framework for representing knowledge. The psychology of computer vision. P. Winston. New York, Mc Graw Hill.
- Mittal, V. O. and C. L. Paris (1993). Context: Identifying its elements from the communication point of view. Workshop on "Using knowledge in its Context", IJCAI-93, Chambéry.
- Montabord, E. (1996). Base de connaissances, hypermédia et contexte, Rennes 1.
- Montabord, E., S. Garlatti, et al. (1994). Interaction between a hypermedia and a knowledge base management system about the brain. neural networks and Expert Systems in Medecine and Healthcare, Plymouth, University of Plymouth.
- Montabord, E., B. Gibaud, et al. (1995). HYPER-YAKA : base de connaissances et hypermédia sensible au contexte. Langage et Modèles à Objets, Nancy.
- Montabord, E., B. Gibaud, et al. (1993). An hypermedia system to manage anatomical knowledge about brain. Computer Assisted Radiology, Berlin, Springer-Verlag.

- Nanard, J. and M. Nanard (1995). "Hypertext Design Environments and the Hypertext Design Process." Communication of the ACM 38(8): 49-56.
- Nanard, M. and J. Nanard (1989). Macweb: un outil pour élaborer des documents. Workshop on Object Oriented Document Manipulation, WOOD-MAN'89.
- Nanard, M. and J. Nanard (2004). Vers une spécification par intention de la production de documents. Intelligence collective : partage et redistribution des savoirs,, Nîmes.
- Nanard, M. and J. Nanard (2005). Saphir : un cadre de référence pour une spécification des hypermédias par intention. H2PTM, Hypermédias Hypertexts Products Tools and Methods, Lavoisier.
- Natarajan, K., M. G. Cawley, et al. (1991). "A knowledge based system paradigm for automatic interpretation of CT scan." Medical informatics 16(2): 167-181.
- Nguyen, G. T. and D. Rieu (1991). Database Issues in Object Oriented Design. Tools'91, Paris.
- Niggemann, J. (1990). "Analysis and Representation of Anatomical Knowledge." Applied Art. Intell.(4): 309-336.
- Oberlander, J. and M. Milasavljevic (1997). Flexible Hypertext - introduction. Hypertext'97, Proceedings of Flexible Hypertext Workshop.
- Ono, M., S. Kubik, et al. (1991). Atlas of the cerebral sulci. Stuttgart, Georg Thieme Verlag.
- Paris, C., M. F. Wu, et al. (2004). Myriad: An Architecture for Contextualized Information Retrieval and Delivery. Proceedings of the 3th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems. Eindhoven, The Netherlands, Springer: 205-214.
- Perfetti, C. A., J.-F. Rouet, et al. (1999). Toward a theory of documents representation. The construction of mental representations during reading. H. v. Oostendorp and S. R. Goldman. Mahwah, NJ, Lawrence Erlbaum Associates: 99-122.
- Pernin, J.-P. and A. Lejeune (2004). Dispositifs d'apprentissage instrumentés par les technologies : vers une ingénierie centrée sur les scénarios. actes du colloque TICE 2004, Compiègne.
- Perrenoud, P. (1998). "La transposition didactique à partir de pratiques : des savoirs aux compétences." Revue des sciences de l'éducation, Montréal XXIV(3): 487-514.
- Petrelli, D., A. D. Angeli, et al. (1999). A User-centered approach to user modeling. UM99, Proceedings of the Seventh International Conference. J. Kay, Springer: 255-264.
- Pohl, W. (1996). Learning about the user. User modeling and machine learning, Workshop: Machine learning meets Human-Computer Interaction, ICML'96. V. Moustakis and J. Herrmann: 29-40.
- Pottier, B. (1974). Linguistique générale, théorie et description, Klincksieck.
- Prié, Y. and S. Garlatti (2005). "Métadonnées et annotations dans le sémantique web." Revue Information - Interaction - Intelligence, Cépaduès, Toulouse Hors série 2004.
- Randell, D. A., Z. Cui, et al. (1992). A Spatial Logic based on Regions and Connection. Int. Conf on Knowledge Representation and Reasoning, Boston.
- Ranwez, S. (2000). Composition Automatique de Documents Hypermédia Adaptatifs à partir d'Ontologies et de Requêtes Intentionnelles de l'Utilisateur. Montpellier, Université de Montpellier 2.
- Ranwez, S. and M. Crampes (1999). Conceptual document and hypertext documents are two different forms of virtual document. Workshop on Virtual Document, Hypertext Functionality and the Web, Toronto.

- Rastier, F. (1997). "Représentation du contenu lexical." Langages, sémantique et intelligence artificielle(87).
- Rekik, Y. A., C. Vanoirbeek, et al. (1999). Reusing Dynamic Document Fragment through Virtual Documents: Key Issues In Document Engineering. Workshop sur les Documents Virtuels Personnalises, IHM'99, Montpellier.
- Rich, E. (1989). Stereotypes and user modeling. user models in dialog systems. A. Kobsa and W. Wahlster. Berlin, Springer Verlag: 35-51.
- Robinson, G. P., A. C. F. Colchester, et al. (1993). Model-Based Recognition of Anatomical Objects from Medical Images. Lecture Notes in Comp. Science: Information Proc. in Med. Imag. H. Barrett and A. Gmitro, Springer-Verlag. 687: 197-211.
- Rosis, F. D., B. d. Carolis, et al. (1993). User Tailored Hypermedia Explanations. INTERCHI'93 Adjunct Proceedings, Amsterdam.
- Roth, E. M. and E. J. Shoben (1983). "The effect of context on structure of categories." Cognitive Psychology(15): 346-378.
- Rouet, F.-F., M. Favard, et al. (1997). "Studying and using multiple documents in History: Effects of discipline expertise." Cognition and Instruction 15(1): 85-106.
- Schaltenbrand, G. and W. Wahren (1977). Atlas for stereotaxy of the human brain. Stuttgart, Georg Thieme Verlag.
- Schank, R. C. and R. P. Abelson (1997). Scripts, plans, goals and understanding, Hillsdale, NJ : Erlbaum.
- Sebillotte, S. (1995). "Methodology guide to task analysis with the goal of extracting relevant characteristics for human computer interfaces." International journal of Human computer interaction 7(4): 341-363.
- Sebillotte, S. and D. L. Scapin (1994). "From users' task knowledge to high-level interface specification." International journal of Human computer interaction 6(1): 1-15.
- Seitz, R. J., C. Bohm, et al. (1990). "Accuracy and precision of the computerized brain atlas programme for localization and quantification in Positron Emission Tomography." Journal of cerebral blood flow and metabolism(10): 443-457.
- Sensevy, G., A. Mercier, et al. (2000). "Vers un modèle de l'action didactique du professeur. A propos de la course à 20." La Pensée sauvage, Grenoble, Recherches en didactique des mathématiques 20(3).
- Sharples, M., B. du Boulay, et al. (1995). The MR Tutor: Computer-based Training and Professional Practice. World Conference on Artificial Intelligence and Education (AI-ED '95), Washington DC.
- Sharples, M. and S. Garlatti (1996). The Use of Computerised Brain Atlas to Support Case-Based Radiology Training Systems, Télécom Bretagne.
- Simmons, R. and R. Davis (1993). The Role of Knowledge and Representation in Problem Solving. Second Generation Expert Systems. J. M. David, J. P. Krivine and R. Simmons, Springer Verlag: 27-45.
- Sokolowska, E. and J. A. Newell (1986). "Multi-Layered Image Representation: Structure and Application in Recognition of Parts of Brain Anatomy." Pattern Recognition Letters(4): 223-230.
- Staemmler, M., E. Claridge, et al. (1993). "SAMMIE - Software applied to multimodal images and education". IMAC 93, Berlin.
- Stevenson, I. (2003). Pedagogical Model for Candle. Technical Report deliverable WP6 of IST CANDLE Project.

- Stroustrup, B. (1986). Multiple inheritance for C++. EUUG.
- Szikla, G., G. Bouvier, et al. (1977). Angiography of the human cortex. Berlin, Heidelberg, New York, Springer Verlag.
- Talairach, J. and P. Tournoux (1988). Co-Planar Stereotactic Atlas Of The Human Brain. Stuttgart, Georg Thieme Verlag.
- Tazi, S. and Y. Altawki (1999). Création de documents virtuels - Cas des supports de cours. Workshop sur les Documents Virtuels Personnalises, IHM'99,, Montpellier.
- Tchounikine, P. (2002). "Pour une ingénierie des Environnements Informatiques pour l'Apprentissage Humain." Revue I3 2(1): 59-95.
- Tchounikine, P. and Al. (2004). Platon-1: quelques dimensions pour l'analyse des travaux de recherche en conception d'EIAH. Rapport de l'action spécifique "Fondements théoriques et méthodologiques de la conception des EIAH", Département STIC, CNRS.
- Thüring, M., J. Hannemann, et al. (1995). "Hypermedia and Cognition: Designing for Comprehension." Communication of the ACM 38(8): 57-66.
- Toga, A. W. (1991). "A Three-Dimensional atlas of structure/function relationships." Journal of chemical neuroanatomy(4): 313-318.
- Towle, A. (1991). Critical Thinking: The Future of Undergraduate Medical Education. London, King's Fund Centre.
- Trichet, F. (1998). DSTM: un environnement de modélisation et d'opérationnalisation de la démarche de résolution de problèmes d'un Système à Base de Connaissances. Nantes, Université de Nantes.
- Tricot, A. and J. Nanard (1997). Un point sur la modélisation des tâches de recherche d'informations dans le domaine des hypermédias. Hypertextes et Hypermédia. J. P. Balpe, A. Lelu, M. Nanrd and I. Saleh, Hermes: 35-56.
- Tricot, A. and F. Plegat-Soutjis (2003). "Pour une approche ergonomique de la conception d'un dispositif de formation à distance utilisant les TIC." Revue STICEF(10).
- Tyler, S. W. and S. Treu (1989). "An interface architecture to provide Adaptive Task-Specific Context for the user." International Journal of Man, Machine and Studies(30): 303-327.
- Vassileva, J. (1996). "A Task-Centered Approach for user Modeling in a hypermedia Office Documentation system." User Models and User Adapted Interaction(6): 185-223.
- Vercoustre, A.-M., J. Dell'Oro, et al. (1997). Reuse of Information through virtual documents. Second Australian Document Computing Symposium, Melbourne, Australia.
- Waern, Y. (1986). Cognitive Aspects of Computer Supported tasks. New York, Wiley.
- Watters, C. and M. Shepherd (1999). Research issues for virtual documents. Workshop on Virtual Document, Hypertext Functionality and the Web, Toronto.
- Wenger, E. (1998). Communities of Practice - Learning, Meaning and Identity. Cambridge, Cambridge University Press.
- Wielinga, B., W. V. d. Velde, et al. (1992). The KADS Knowledge Modelling Approach. Proceedings of the 2nd Japanese Knowledge Acquisition for Knowledge-Based Systems Workshop. R. Mizoguchi and H. Motoda. Hitachi, Advanced Research Laboratory, Hatoyama, Saitama, Japan: 23-42.
- Willamowski, J. (1992). Modélisation de tâches pour la résolution de problèmes en coopération système-utilisateur. Grenoble, Université Joseph Fourier.

- Winston, M. E., R. Chaffin, et al. (1987). "A Taxonomy of Part-Whole Relations." Cognitive Science(11): 417-444.
- Wolinski, F. and F. Perrot (1991). Representation of Complex Objects: Multiple Facts with Part-Whole Hierarchies. ECOOP.
- Woods, W. A. (1975). What's in a link: Foundations for Semantics Networks. Representation and Understanding: Studies in Cognitive Science. D. G. Bobrow and A. Collins. New York, Academic Press.
- Zachmann, H. H. (1991). Interpretation of cranial MR images using a digital atlas of the human head. CAR 91, Berlin, Springer Verlag.

